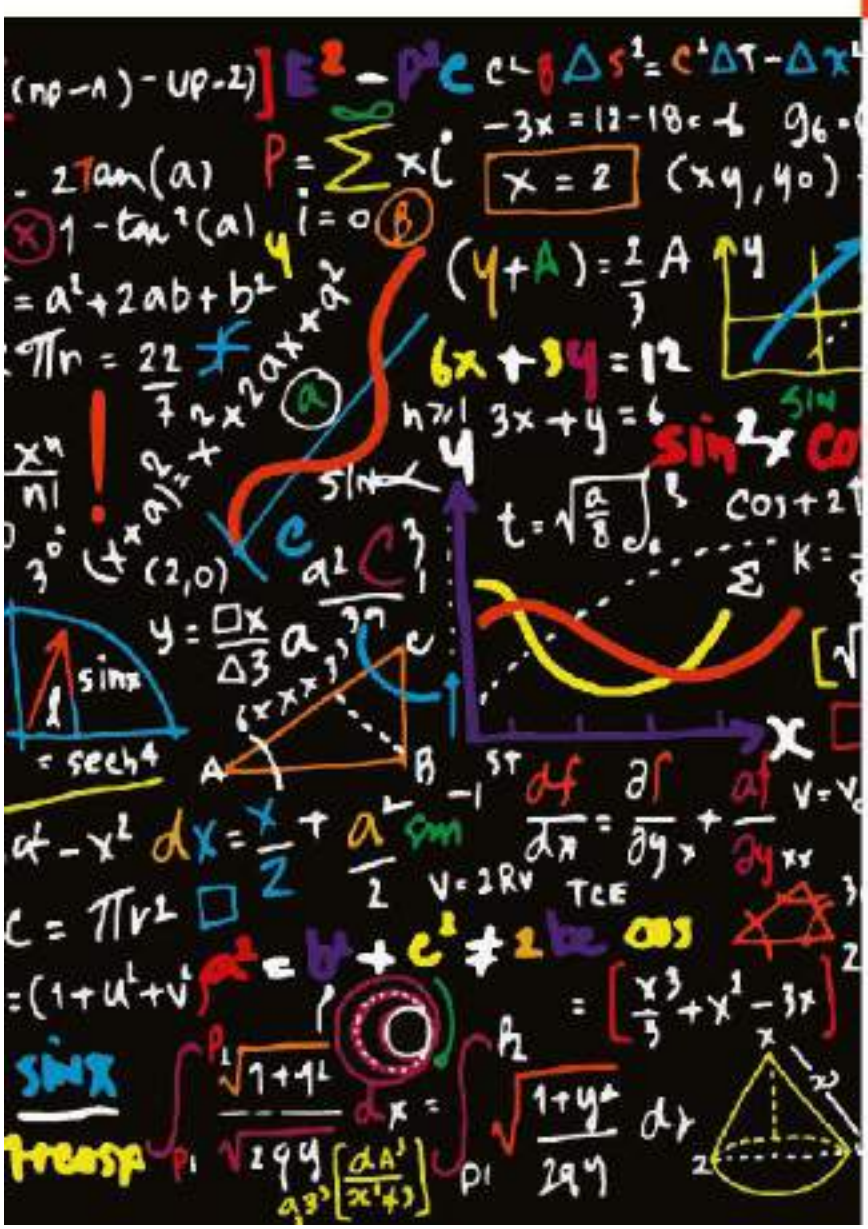
ALESSANDRA BERNARDI,
ALESSANDRO GIMIGLIANO

Algebra lineare e geometria analitica

Seconda edizione



CittàStudi
EDIZIONI

Alessandra Bernardi, Alessandro Gimigliano

Algebra lineare e geometria analitica

Seconda edizione

CittàStudi
EDIZIONI

CittàStudi
EDIZIONI

www.cittastudi.it

Proprietà letteraria riservata
© 2018 De Agostini Scuola SpA – Novara
2ª edizione: maggio 2018
Printed in Italy

In copertina : iStock
Grafico: Marco Fiorello, Matteo Rossi
Art Director: Nadia Maestri

Tutti i diritti riservati. Nessuna parte del materiale protetto da questo copyright potrà essere riprodotta in alcuna forma senza l'autorizzazione scritta dell'Editore.

Fotocopie per uso personale del lettore possono essere effettuate nei limiti del 15% di ciascun volume/fascicolo di periodico dietro pagamento alla SIAE del compenso previsto dall'art. 68, comma 4, della legge 22 aprile 1941 n. 633.
Le riproduzioni ad uso differente da quello personale potranno avvenire, per un numero di pagine non superiore al 15% del presente volume/fascicolo, solo a seguito di specifica autorizzazione rilasciata da CLEARedi, Centro Licenze e Autorizzazioni per le Riproduzioni Editoriali, Corso di Porta Romana 108, 20122 Milano – e-mail: *autorizzazioni@clearedi.org*.

Stampa: Stampatre - Torino

Ristampe:	0 1	2 3	4 5	6 7	8 9
Anno:	2018	2019	2020	2021	2022

Indice

0 Introduzione **5**
- 0.1 Per gli studenti 5
- 0.2 Prerequisiti 6
- 0.3 Per i docenti 8
- 0.4 Rispetto alla prima edizione 9

1 Elementi di teoria degli insiemi **11**
- 1.1 Definire un insieme 11
- 1.2 Operazioni fra insiemi 14
- 1.3 Insiemi numerici, principio d'induzione 15
- 1.4 Funzioni 17
- 1.5 Operazioni, strutture algebriche 20
- 1.6 Esercizi 23

2 Spazi vettoriali **25**
- 2.1 Introduzione: vettori geometrici 25
- 2.2 Spazi vettoriali 29
- 2.3 Indipendenza lineare, basi 31
- 2.4 Sottospazi 39
- 2.5 Esercizi 43

3 Le matrici **49**
- 3.1 Introduzione 49
- 3.2 Generalità sulle matrici 49
- 3.3 Il prodotto fra matrici 56
- 3.4 L'anello delle matrici quadrate 61
- 3.5 Esercizi 64

4 Sistemi lineari **69**
- 4.1 Introduzione 69
- 4.2 Risoluzione di sistemi lineari 72
- 4.3 Algoritmo per l'eliminazione di Gauss 79
- 4.4 Sistemi omogenei 83
- 4.5 Rango di una matrice 84
- 4.6 Esercizi 88

5 Applicazioni lineari **95**
5.1 Introduzione 95
5.2 Applicazioni lineari fra spazi vettoriali 96
5.3 Nucleo e immagine, iniettività e suriettività 99
5.4 Applicazioni lineari e matrici 106
5.5 Cambiamenti di base, matrici invertibili 112
5.6 Esercizi 117

6 Il determinante **125**
6.1 Introduzione 125
6.2 Il determinante 126
6.3 Proprietà del determinante 132
6.4 Determinante e permutazioni 136
6.5 Applicazioni del determinante 138
6.5.1 Matrici che dipendono da parametri 138
6.5.2 Matrici non quadrate 139
6.5.3 Calcolo dell'inversa 141
6.5.4 Regola di Cramer 143
6.5.5 Prodotto vettoriale 143
6.6 Esercizi 144

7 Autovalori e autovettori **149**
7.1 Introduzione 149
7.2 Autovalori, autovettori, autospazi 150
7.3 Ricerca degli autovalori 154
7.4 Endomorfismi semplici, diagonalizzazione 160
7.5 Esercizi 166

8 Spazi vettoriali euclidei **171**
8.1 Prodotto scalare e spazi vettoriali euclidei 171
8.2 Norma, ortogonalità 175
8.3 Basi ortonormali 180
8.4 Isometrie 184
8.5 Il teorema spettrale 187
8.6 Esercizi 191

9 Geometria analitica nello spazio **195**
9.1 Introduzione 195
9.1.1 Premesse 195
9.2 Piani nello spazio 198
9.2.1 Equazione cartesiana ed equazioni parametriche del piano 198
9.3 Rette nello spazio 204
9.3.1 La retta, equazioni cartesiane e parametriche . . . 204
9.3.2 Retta per due punti dati 207
9.3.3 Calcolo dei parametri direttori di una retta 209
9.3.4 Passaggio da equazioni parametriche a cartesiane e viceversa 210

9.3.5 Posizioni reciproche di due rette 211
9.3.6 Posizioni reciproche di retta e piano 213
9.4 Il prodotto vettoriale 220
9.5 Esercizi 223

10 Forme bilineari e quadratiche **229**
10.1 Forme bilineari 229
10.2 Forme quadratiche 236
10.3 Coniche 243
10.3.1 Introduzione 243
10.3.2 Le coniche come luoghi geometrici 245
10.3.3 Classificazione 249
10.3.4 Coniche e isometrie 252
10.4 Quadriche 255
10.4.1 Classificazione delle quadriche 256
10.5 Esercizi 258

11 Spazi affini **261**
11.1 Gli spazi affini 261
11.2 Sottospazi affini 263
11.3 Affinità 267
11.4 Esercizi 268

12 Numeri complessi **271**
12.1 Introduzione 271
12.2 Rappresentazione algebrica 272
12.3 Rappresentazione cartesiana 275
12.4 Rappresentazione trigonometrica 277
12.5 Rappresentazione esponenziale 282
12.6 Esercizi 285

13 Soluzione di alcuni esercizi **287**
13.1 Soluzioni Capitolo 1 287
13.2 Soluzioni Capitolo 2 288
13.3 Soluzioni Capitolo 3 290
13.4 Soluzioni Capitolo 4 292
13.5 Soluzioni Capitolo 5 296
13.6 Soluzioni Capitolo 6 301
13.7 Soluzioni Capitolo 7 303
13.8 Soluzioni Capitolo 8 307
13.9 Soluzioni Capitolo 9 308
13.10 Soluzioni Capitolo 10 311
13.11 Soluzioni Capitolo 11 312

Bibliografia **313**

Indice analitico **315**

L'Editore ringrazia

Amedeo Altavilla *Università degli Studi di Roma "Tor Vergata"*
Cristiano Bocci, *Università degli Studi di Siena*
Maria Chiara Brambilla, *Università Politecnica delle Marche*
Stefano Capparelli, *Università degli Studi di Roma "La Sapienza"*
Alessandra Cherubini, *Politecnico di Milano*
Michele Di Cristo, *Politecnico di Milano*
Giovanni Luca Dierna, *Università degli Studi di Palermo*
Paolo Dulio, *Politecnico di Milano*
Lara Ercoli, *Università degli Studi di Brescia*
Patrizio Frosini, *Università di Bologna*
Vincenzo Giordano, *Politecnico di Bari*
Remke Nanne Kloosterman, *Università degli Studi di Padova*
Vincenzo La Rosa, *Università degli Studi di Palermo*
Claudia Landi, *Università degli Studi di Modena e Reggio Emilia*
Giovanni Manno, *Politecnico di Torino*
Valentina Pepe, *Università degli Studi di Roma "La Sapienza"*
Emanuele Rodaro, *Politecnico di Milano*
Stefano Siboni, *Università degli Studi di Trento*
Luis Eduardo Solá Conde, *Università degli Studi di Trento*
Paola Lea Staglianò, *Università degli Studi di Palermo*
Simone Ugolini, *Università degli Studi di Trento*
Mauro Zannetti, *Università degli Studi dell'Aquila*
Andrea Zavaglia, *Politecnico di Milano*

I loro preziosi suggerimenti hanno contribuito alla realizzazione di questa edizione di *Algebra lineare e geometria analitica*

CAPITOLO 0

Introduzione

Diamo qui alcune avvertenze per l'uso del libro soprattutto agli studenti, ma anche ai docenti che lo adotteranno.

0.1 Per gli studenti

Benvenuti al vostro primo corso di *geometria e algebra lineare*. Gli argomenti sviluppati in questa parte della matematica sono in gran parte del tutto nuovi rispetto al vostro percorso nella scuola media superiore, anche se non mancano gli agganci con esso, per esempio:

- studio di sistemi lineari di equazioni;
- uso della geometria analitica (cartesiana) per descrivere oggetti geometrici (per esempio rette o circonferenze nel piano) con strumenti algebrici (equazioni);
- uso dei vettori (incontrato anche in fisica).

Riguardo al primo punto, studieremo metodi generali per la soluzione di sistemi lineari di equazioni (cioè sistemi dove le equazioni sono date da polinomi di primo grado nelle variabili), e per determinare (senza risolverli) se essi abbiano soluzioni o no, e "quante" ne abbiano.

La geometria analitica che studieremo sarà incentrata nel caso dello spazio tridimensionale (in genere non trattato alle superiori) e riguarderà soprattutto rette e piani nello spazio.

I vettori saranno in un certo senso alla base del corso, in quanto gran parte di esso studierà strutture (spazi vettoriali) che generalizzano il concetto di vettore geometrico per usarne le idee base anche in altri contesti.

Il testo è ricco di esempi ed esercizi che seguono la definizione delle varie strutture che vengono introdotte e delle loro proprietà, cosicché possano aiutare a prendere familiarità con esse e a riuscire a risolvere

problemi che le riguardano. Gli esercizi più complessi sono segnalati da un asterisco.

Nello studio di questa materia (come di molte altre) è importante riuscire ad armonizzare i due aspetti che caratterizzano la padronanza dei concetti: il comprendere la definizione delle strutture studiate e delle loro proprietà (definizioni, teoremi) e saperle trattare da un punto di vista più pratico (saper vedere dove e come applicarle, risolvere problemi ed esercizi).

Una preparazione troppo "teorica" rischia di portare a imparare (magari a memoria) le nozioni fondamentali del corso e le dimostrazioni dei teoremi, ma non a saper poi svolgere semplici esercizi dove si richiede di utilizzare tali nozioni. D'altro canto una preparazione tutta "per esercizi" rischia di portare a saper risolvere varie tipologie di esercizi apprendendo il tipo di procedimento richiesto (i "conti che bisogna fare"), senza poi saper rispondere a semplici domande teoriche (se non si è mai riflettuto sul significato delle definizioni), perdendo un po' il senso di quello che stiamo facendo. Un buon esercizio di autovalutazione è quello di cercare di inventarsi esempi delle varie strutture cercate, come:

- fare un esempio di uno spazio vettoriale e di un suo sottospazio non banale;
- fare un esempio di uno spazio vettoriale e di una sua base (escludendo le basi canoniche in $\mathbb{R}^n$);
- fare un esempio di un'applicazione lineare da $\mathbb{R}^2$ a $\mathbb{R}^3$;
- fare un esempio di un'applicazione lineare da $\mathbb{R}^3$ a $\mathbb{R}^3$ che non sia iniettiva;
- fare un esempio di una matrice 4×3 di rango 2.

Saper creare da soli esempi dei vari concetti che si sono studiati è un ottimo segno del fatto che si sono veramente compresi.

Per quanto riguarda le dimostrazioni dei teoremi, essere in grado di seguirle e averle capite non è utile solo nozionisticamente, ma vuol anche dire aver assimilato meglio i concetti che stiamo studiando. Abbiamo inserito alcune delle dimostrazioni in caratteri più piccoli quando, a nostro avviso, sono meno essenziali o più complicate.

0.2 Prerequisiti

I prerequisiti specifici per questo corso non sono molti, e dovrebbero far parte delle conoscenze di chiunque abbia superato un ciclo di scuola media superiore: conoscenza del calcolo letterale (espressioni letterali, polinomi notevoli), risoluzione di equazioni, elementi di geometria analitica nel piano, elementi di geometria euclidea e trigonometria.

Vediamo in qualche dettaglio alcune delle conoscenze e capacità di base che si dovrebbero avere:

Calcolo letterale

È necessario aver familiarità con le nozioni di base del calcolo letterale, per esempio essere capaci di distinguere fra espressioni di tipo $ax-3$ e $a(x-3)$, capaci di semplificare correttamente $\frac{3x+9}{3y}$, $\frac{x+xy}{x}$ e $\frac{x^2+3x}{x+3}$ e saper dire sotto quali condizioni esse siano definite o meno. Bisogna conoscere i polinomi notevoli elementari:

$$(a+b)^2 = a^2+2ab+b^2;\ (a+b)^3 = a^3+3a^2b+3ab^2+b^3;\ (a+b)(a-b) = a^2-b^2,$$

e saper riconoscere queste forme anche in espressioni leggermente più complicate, per esempio:

$$\frac{4a^4 - x^8y^4}{\sqrt{2}a + x^2y} = \frac{(2a^2 + x^4y^2)(2a^2 - x^4y^2)}{\sqrt{2}a + x^2y} =$$

$$= \frac{(2a^2 + x^4y^2)(\sqrt{2}a + x^2y)(\sqrt{2}a - x^2y)}{\sqrt{2}a + x^2y} = (2a^2 + x^4y^2)(\sqrt{2}a - x^2y),$$

notando che tutta l'espressione è definita solo se $\sqrt{2}a + x^2y \neq 0$.

Risoluzione di equazioni

Bisogna saper risolvere equazioni polinomiali di primo e secondo grado in presenza di parametri, cioè, saper valutare che cosa succede al variare dei parametri a, b, c in equazioni come $ax + b = 0$ e $ax^2 + bx + c = 0$. Ad esempio $ax + b = 0$ non ha soluzioni (è *impossibile*) se $a = 0$ e $b \neq 0$, mentre ne ha infinite (è *indeterminata*) se $(a, b) = (0, 0)$.

Esempio 0.2.1. Se vogliamo discutere che cosa accada nell'equazione $k^2x + kx = k$, al variare del parametro k, avremo:

$$k^2x + kx = k \quad \leftrightarrow \quad (k^2 + k)x - k = 0;$$

quindi se $k^2 + k \neq 0$ (cioè $k \neq 0$ e $k \neq -1$), si avrà un'unica soluzione:

$$x = \frac{k}{k^2 + k}.$$

Se invece $k = -1$, l'equazione diviene $0x = -1$, cioè $0 = -1$ che è impossibile.

Se infine $k = 0$, l'equazione diviene $0x = 0$, cioè $0 = 0$ che è un'identità (sempre vera), quindi ogni valore di x è soluzione e l'equazione è indeterminata.

L'equazione $ax^2 + bx + c = 0$ ha formula risolutiva (per $a \neq 0$):

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

Il termine $\Delta = \sqrt{b^2 - 4ac}$ si dice *discriminante* dell'equazione, e se $\Delta < 0$ essa non ha soluzioni reali, ne ha una sola (si dice anche che ne ha due coincidenti) se $\Delta = 0$ e infine ne ha due distinte quando $\Delta > 0$.

Esempio 0.2.2. Dire che cosa accade, al variare di k, alle radici dell'equazione $x^2 + 2kx + k^2 - k + 1 = 0$.

Si ha che $\Delta = 4k^2 - 4(k^2 - k + 2) = 4k - 8$. Quindi per $k < 2$ l'equazione non ha soluzioni ($\Delta < 0$), ne ha due distinte per $k > 2$, e cioè:

$$x = \frac{2k \pm \sqrt{4k-8}}{2} = k \pm \sqrt{k-2},$$

mentre per $k = 2$ l'unica soluzione è $x = 2$.

È importante non farsi sviare da equazioni polinomiali (di qualsiasi grado) che appaiono in forma fattorizzata; per esempio per risolvere l'equazione: $(x - 4)(x + 3)(x - 6) = 0$, non bisogna effettuare la moltiplicazione, ottenendo:

$$(x-4)(x+3)(x-6) = x^3 - 6x^2 - 7x + 72 = 0,$$

che risulta una complicata equazione di terzo grado; in effetti la forma $(x-4)(x+3)(x-6) = 0$ mostra già che le soluzioni sono $x_1 = 4$, $x_2 = -3$ e $x_3 = 6$, in quanto un prodotto si annulla se e solo se si annulla almeno uno dei suoi fattori.

Elementi di geometria analitica nel piano

Vanno conosciuti i principi base dell'utilizzo delle coordinate cartesiane nel piano con la corrispondenza fra punti del piano e coppie ordinate (x, y) di numeri reali, che permette di passare dallo studio di oggetti geometrici a quello di oggetti algebrici e viceversa, come saper trovare la distanza fra due punti:

$$d(A, B) = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2},$$

o per esempio saper studiare le rette, rappresentate da equazioni del tipo $ax + by + c = 0$ (con $(a, b) \neq (0, 0)$), e tali che ogni equazione di questo tipo rappresenta una retta. Collegata a tale rappresentazione è la nozione di *coefficiente angolare*, che ci permette di riconoscere le condizioni di parallelismo e perpendicolarità fra rette.

Il nostro studio della geometria analitica non verterà su questi argomenti (geometria analitica nel piano), ma soprattutto sullo studio della geometria analitica nello spazio tridimensionale, che ne è una naturale estensione. Le nozioni della geometria nel piano dovrebbero aiutare, cogliendo analogie e differenze, nello studio di rette e piani nello spazio.

0.3 Per i docenti

L'idea base di questo libro è di avere un testo di riferimento per un corso base di algebra lineare e geometria analitica, come si ha per esempio in gran parte dei Corsi di Laurea in Ingegneria. Il testo non è invece strutturato per costituire un testo di riferimento generale per queste materie,

non avendo il carattere di completezza e di ricchezza di argomenti che un testo di riferimento richiederebbe, anche quando questi argomenti non vengono svolti nel corso che lo adotta (alcuni testi di riferimento sono indicati in bibliografia).

Non si introducono, ad esempio, gli spazi proiettivi, né la forma canonica di Jordan o le curve algebriche; gli spazi vettoriali su cui si lavora sono sempre e solo sui reali, non si trattano forme hermitiane sui complessi.

L'idea è quella di avere un testo abbastanza agile e "student friendly"; ad esempio non mancano risultati che restano solo enunciati e non dimostrati (anche se non molti) e alcune dimostrazioni più complesse sono scritte a caratteri piccoli perché sono presenti per completezza nel testo, ma spesso non richieste nel programma d'esame.

Nel suddividere gli argomenti nei vari capitoli si è cercato di rendere tali capitoli il più possibile autonomi, in modo da permettere facilmente uno svolgimento del corso che segua un ordine diverso nella presentazione degli argomenti rispetto all'ordine dei capitoli del libro; per esempio si può preferire un approccio meno "teorico" e svolgere prima gli argomenti *Matrici* (Cap. 3) e *Sistemi lineari* (Cap.4) rispetto agli *Spazi vettoriali* (Cap.2); oppure si può preferire di anticipare la parte di *Geometria analitica* (Cap. 9) a subito dopo i sistemi lineari.

Alcuni argomenti (forme bilineari e quadratiche, spazi affini, numeri complessi), che costituiscono gli ultimi capitoli del libro, sono in genere non compresi nei corsi usuali per ingegneri, ma sono stati inseriti in quanto alcuni docenti preferiscono svolgerli. In particolare:

- il Capitolo 10 tratta le forme bilineari e quadratiche e contiene anche la trattazione delle coniche e delle quadriche; quest'ultima (che inizia con la sezione 10.3) è sviluppata in modo da poter esser trattata anche senza introdurre le forme bilineari, iniziando direttamente da quella sezione.

- il Capitolo 11 tratta gli spazi affini. Può essere premesso al Capitolo 9 se si vuole trattare la geometria analitica nello spazio usando questo linguaggio.

Si è cercato poi di avere un buon numero di esempi ed esercizi, sia frammisti all'esposizione della teoria, sia situati alla fine di ogni capitolo, e si sono inserite (nel Capitolo 13) le soluzioni per una selezione degli esercizi proposti, mentre le soluzioni di tutti gli altri esercizi sono pubblicate nel sito collegato al libro.

0.4 Rispetto alla prima edizione

I cambiamenti nel presente testo rispetto alla prima edizione sono essenzialmente i seguenti:

- Sono stati aggiunti nuovi esercizi e introdotti gli asterischi che segnalano gli esercizi più complessi.
- Capitolo 4: revisione e dimostrazione dell'uguaglianza di rango per righe e per colonne di una matrice.
- Capitolo 6: revisione della dimostrazione che $\det(A) = \det({}^tA)$.
- Capitolo 8: si è scelto la denominazione *spazi vettoriali euclidei* per spazi vettoriali dotati di prodotto scalare, in quanto la denominazione *spazi euclidei* è più generalmente usata per spazi affini dotati di prodotto scalare.
- Capitolo 10: Nasce come ampliamento dell'Appendice B della prima edizione (*Coniche e Quadriche*), aggiungendo la trattazione di forme bilineari e quadratiche.
- Capitolo 11: *Spazi Affini*, assente nell'edizione precedente.

CAPITOLO 1

Elementi di teoria degli insiemi

Prerequisiti: Nessuno.

Obiettivi: Conoscere il linguaggio della teoria degli insiemi e la sua simbologia; aver chiari i concetti di funzione, funzione iniettiva, funzione suriettiva; conoscere i principali insiemi numerici e le più semplici strutture algebriche.

1.1 Definire un insieme

La costruzione di una rigorosa (assiomatica) teoria degli insiemi è una procedura molto complessa, che ha richiesto un grande lavoro in logica matematica agli inizi del secolo scorso, ed è fuori dalla portata e dallo scopo di questo testo; quella che presenteremo qui è una breve introduzione agli insiemi e alle notazioni e operazioni che si usano con essi. Si tratta di essere sicuri di possedere le nozioni fondamentali per utilizzare il linguaggio (e la simbologia) di questa teoria, che è ormai uno strumento essenziale per esprimere concetti e comunicare in matematica.

Prima di iniziare con la teoria degli insiemi vera e propria, premettiamo la definizione di alcuni simboli logici, di uso comune in matematica:

- $\forall$ = "per ogni";
- $\exists$ = "esiste";
- $\exists!$ = "esiste un unico".

Il loro uso tipico è in espressioni come le seguenti:

$$\forall x \geq 0,\ \exists y \text{ tale che } y^2 = x.$$

Che si legge: "per ogni x maggiore o uguale a zero, esiste un y tale che $y^2 = x$" (ad esempio tale enunciato è vero nell'insieme dei numeri reali, falso nell'insieme dei numeri interi).

$$\forall x,\ \exists!\ y \text{ tale che } y + x = 0.$$

Che si legge: "per ogni x, esiste un unico y tale che $y + x = 0$".
Esercizio: Tale enunciato è vero nell'insieme dei numeri interi?

Altri simboli logici sono $\Rightarrow$ (implica) e $\Leftrightarrow$ (coimplica); ad esempio:

$$x \geq 0 \Leftrightarrow \exists\, y, y^2 = x.$$

Che si legge: "x è maggiore o uguale a zero se e solo se esiste un y tale che $y^2 = 0$" (tale enunciato è vero nell'insieme dei numeri reali).

L'idea intuitiva di un **insieme** è quella di una "collezione di oggetti" (quelli che chiameremo elementi dell'insieme), ad esempio: "l'insieme di tutti gli studenti dell'Ateneo bolognese", "l'insieme di tutte le auto francesi", "l'insieme dei numeri interi dispari".

Un insieme resta definito ogniqualvolta sia fissato un "Universo di discorso" U (gli oggetti di cui si parla, ad esempio "i numeri", "i poligoni", "gli esseri umani"), nel quale sceglieremo gli elementi dell'insieme, e una proprietà che ci permetterà di dire di ogni elemento di U se appartiene o no all'insieme. Ad esempio, se U è dato da "i poligoni", la proprietà "avere quattro lati" individua l'insieme formato da tutti i quadrilateri, che si scriverà così (gli insiemi si denoteranno in genere con lettere maiuscole, gli elementi con lettere minuscole):

$$A = \{x \mid x \text{ sia un poligono di quattro lati}\}.$$

Altro esempio: se considiamo la frase "x è un numero intero positivo e minore di 13" (qui U sarà dato, ad esempio, da tutti i numeri interi), allora tale proprietà definisce l'insieme costituito dagli elementi $1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$.

Ci sono due modi di assegnare un insieme:

1. dare l'elenco completo dei suoi elementi;
2. definirlo tramite una proprietà che i suoi elementi devono soddisfare.

Per la scrittura in simboli di un insieme A si adottano le parentesi graffe, vediamo alcuni esempi:
Definiti nel modo 1:

- $A = \{1, 3, 5, 6\}$;
- $B = \{\diamondsuit, \clubsuit, \heartsuit, \spadesuit\}$;
- $C = \{a, b, c, d, e\}$.

Definiti nel modo 2:

- $D = \{x \mid x \text{ sia un carciofo}\}$;
- $E = \{x \mid x \text{ sia un numero primo}\}$;
- $F = \{x \mid x \text{ sia un intero e } 3 \leq x < 13\}$.

Nei primi tre esempi è fornito l'elenco degli elementi dell'insieme; negli altri è definita la proprietà che caratterizza i suoi elementi (ad esempio, nell'insieme D: "essere un carciofo"); la scrittura adottata, e cioè $A = \{x \mid x \text{ sia un carciofo}\}$ si legge: "A è l'insieme degli x tali che x sia un carciofo".

Avremmo potuto anche scrivere l' insieme B così:

$B = \{x \mid x \text{ sia un seme delle carte francesi}\}$ (usando il modo 2),

e invece l'insieme F come

$F = \{3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ (usando il modo 1).

Definizione 1.1.1. Un insieme particolare è quello che non contiene alcun elemento, detto *insieme vuoto*, e indicato con il simbolo $\emptyset$.

Per indicare che un elemento a appartiene all'insieme A si usa il simbolo "$\in$", e si scrive quindi: $a \in A$; ad esempio se A è l'insieme del primo esempio, possiamo scrivere $3 \in A$. Scriviamo invece $15 \notin A$ per indicare che l'elemento 15 non sta in A.

Definizione 1.1.2. Si dice che un insieme B *è contenuto in* A, oppure che B *è un sottoinsieme di* A, e si scrive: $B \subset A$, quando un insieme B è tale che ogni suo elemento appartiene anche all'insieme A.

Ad esempio, se $A = \{1, 3, 56\}$ e $B = \{1, 3\}$, si ha $B \subset A$.

Si ha che ogni insieme ha almeno due sottoinsiemi: se stesso e l'insieme vuoto (che è quindi sottoinsieme di ogni altro insieme): $\forall A$, $A \subset A$ e $\emptyset \subset A$.

Attenzione:

Se consideriamo gli insiemi $\{1, 3, 4\}$ e $\{1\}$, potremo scrivere $\{1\} \subset \{1, 3, 4\}$, ma NON $\{1\} \in \{1, 3, 4\}$, perché la scrittura $\{1\}$ indica un

insieme che ha solo 1 come suo elemento, quindi esso è un sottoinsieme di $\{1,3,4\}$ e non un suo elemento, perciò va scritto $\{1\} \subset \{1,3,4\}$. Scriveremo invece correttamente $1 \in \{1,3,4\}$.

Notiamo che quando si assegna un insieme, non si considera alcun ordine fra i suoi elementi: $\{1,2,3\}$ e $\{2,3,1\}$ sono due scritture dello stesso insieme. Quando si vuole considerare anche l'ordine degli elementi si usano le parentesi tonde: $(1,2,3)$ e $(2,3,1)$ sono due oggetti diversi (in questo caso si parlerà di *terne ordinate* di elementi, e in modo analogo di coppie ordinate, quaterne ordinate, ecc.).

1.2 Operazioni fra insiemi

Dati due insiemi A, B possiamo considerare l'insieme dato dagli elementi comuni ad A e B, che è detto l'*intersezione* di A e B:

$$A \cap B = \{x \mid x \in A \text{ e } x \in B\}.$$

Ad esempio, se $A = \{1,7,51\}$ e $B = \{1,2,51\}$, avremo:

$$A \cap B = \{1,51\}.$$

Una diversa operazione fra insiemi è invece quella di *unione*:

$$A \cup B = \{x \mid x \in A \text{ oppure } x \in B\}.$$

L'insieme $A \cup B$ è dato da tutti quegli elementi che stanno in almeno uno dei due insiemi.

Ad esempio se $A = \{1,7,51\}$ e $B = \{1,2,51\}$, avremo:

$$A \cup B = \{1,2,7,51\}.$$

Un'ulteriore operazione fra insiemi è invece quella di *sottrazione*:

$$A - B = \{x \mid x \in A \text{ e } x \notin B\}.$$

L'insieme $A - B$ è dato da tutti quegli elementi che stanno in A ma non in B.

Ad esempio se $A = \{1,7,51\}$ e $B = \{1,2,51\}$, avremo:

$$A - B = \{7\}.$$

Un'altra operazione fra due insiemi A e B è quella di *prodotto cartesiano* $A \times B$:

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

Cioè l'insieme $A \times B$ è costituito da tutte le coppie ordinate di elementi, il primo dei quali sta in A e il secondo in B.

Esempio: Siano: $A = \{\clubsuit, \heartsuit\}$ e $B = \{1, 3, 56\}$; allora si avrà:

$$A \times B = \{(\clubsuit, 1), (\clubsuit, 3), (\clubsuit, 56), (\heartsuit, 1), (\heartsuit, 3), (\heartsuit, 56)\}.$$

mentre $A \times A$, che si indicherà anche con A^2, sarà:

$$\{(1,1), (1,3), (1,56), (3,1), (3,3), (3,56), (56,1), (56,3), (56,56)\}.$$

Un altro esempio di prodotto cartesiano è dato dalle coordinate nel piano (dette appunto coordinate cartesiane): se si considera il prodotto cartesiano dell'insieme $\mathbb{R}$ dei numeri reali (vedi più avanti per una sua definizione) per se stesso si otterrà:

$$\mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\} \quad \text{(si indica anche con } \mathbb{R}^2\text{)}.$$

Con tali coppie di coordinate si denotano i punti del piano cartesiano, e da ciò viene il nome di prodotto cartesiano di insiemi, infatti l'insieme $\mathbb{R}^2$ delle coppie di numeri reali si rappresenta (tramite il metodo delle coordinate) con i punti del piano.

1.3 Insiemi numerici, principio d'induzione

Vediamo quali sono i principali insiemi numerici di cui ci serviamo.

- $\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}$ i numeri **naturali**. Si tratta dei numeri più semplici da usare, quelli associati a ogni elementare "conteggio".
- $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, 4, \ldots\}$ i numeri **interi**. Ai numeri naturali abbiamo aggiunto gli interi negativi.
- $\mathbb{Q} = \{\ldots, -2, \ldots, -\frac{3}{2}, \ldots, 0, \ldots, \frac{1}{3}, \ldots, \frac{1}{2}, \ldots, 1, \ldots \frac{4}{3}, \ldots 2, \ldots\}$ i numeri **razionali**. Ai numeri interi abbiamo aggiunto tutte le frazioni.
- $\mathbb{R} = \{\ldots, -\sqrt{2}, -1, 0, 1, \sqrt{2}, \pi, 4, 2\pi, \ldots\}$ i numeri **reali**. Ai numeri razionali abbiamo aggiunto gli irrazionali.

Qualche commento:

- Si ha $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

- In $\mathbb{N}$ per ogni numero esiste un suo *successore*, lo 0 non è successore di nessun numero e numeri diversi hanno successori diversi. Questi, insieme al principio di induzione (vedi più sotto in questa sezione) costituiscono gli assiomi che caratterizzano l'insieme $\mathbb{N}$ (assiomi di Peano), con la sua struttura di catena ordinata infinita con un primo elemento.

- In $\mathbb{Z}$ per ogni numero esiste un suo *successore* e ogni numero è successore di un altro numero. La struttura di ordine di $\mathbb{Z}$ è quella di catena ordinata infinita senza primo elemento. In $\mathbb{Z}$ la sottrazione è sempre eseguibile.

- In $\mathbb{Q}$ non esiste la nozione di successore di un numero; il suo ordine è *denso*, cioè scelti comunque due suoi elementi tra di essi se ne trovano infiniti altri. Si può sempre eseguire la divisione di due numeri in $\mathbb{Q}$, purché il secondo non sia zero. Gli elementi di $\mathbb{Q}$ si possono scrivere in forma decimale, ottenendo sempre decimali finiti o infiniti ma periodici.

- In $\mathbb{R}$ abbiamo tutti i possibili decimali, anche infiniti e non periodici (*irrazionali*, come π o $\sqrt{2}$).

Lavorando con $\mathbb{N}$, c'è un principio molto importante, che si usa in molte dimostrazioni: il *principio di induzione*. Il principio di induzione riguarda la struttura stessa dei numeri naturali (è uno degli assiomi che individuano la struttura di $\mathbb{N}$), e si usa per "dimostrare infinite proprietà in una volta sola".

Principio di induzione Sia $\mathcal{P}$ una proprietà che ha senso per numeri $n \in \mathbb{N}$; allora se valgono:

1) la $\mathcal{P}$ vale per lo 0;

2) quando la $\mathcal{P}$ vale per n, allora vale anche per $n+1$;

si ha che $\mathcal{P}$ vale $\forall\, n \in \mathbb{N}$.

Un altro modo di enunciare lo stesso principio può esprimersi così :

Sia $A \subset \mathbb{N}$; allora se valgono:

1) $0 \in A$;

2) quando $n \in A$, allora anche per $n+1 \in A$;

si ha la $A = \mathbb{N}$.

Basta infatti definire A come l'insieme degli $n \in \mathbb{N}$ per cui vale la $\mathcal{P}$ per passare dalla prima alla seconda formulazione.

Osserviamo che vale anche una forma più generale del principio di induzione:

Siano $A \subset \mathbb{N}$ e $n_0 \in \mathbb{N}$; allora se valgono:

1. $n_0 \in A$;
2. quando $n \in A$, allora anche per $n+1 \in A$;

si ha la $A = \{x \in \mathbb{N} |\ x \geq n_0\}$.

Esempio 1.3.1. Un esempio "classico" di applicazione del principio d'induzione è la dimostrazione che la seguente proprietà $\mathcal{P}(n)$ vale per ogni $n \in \mathbb{N}$:

$$\mathcal{P}(n): \text{ La somma dei naturali da 0 a } n \text{ vale } \frac{n(n+1)}{2}.$$

Dimostriamola usando il principio d'induzione. Innanzi tutto notiamo che vale $\mathcal{P}(0)$, infatti:

$$0 = \frac{0(0+1)}{2}.$$

Vogliamo poi vedere che se supponiamo che valga $\mathcal{P}(n)$, allora vale anche $\mathcal{P}(n+1)$. Consideriamo la somma dei primi naturali, da 0 fino a $n+1$: $0+1+2+\cdots+n+(n+1)$; usando la $\mathcal{P}(n)$ per la somma fino a n avremo:

$$0+1+2+\cdots+n+(n+1) = \frac{n(n+1)}{2} + (n+1) =$$

$$= \frac{n(n+1)+2(n+1)}{2} = \frac{(n+1)(n+2)}{2}$$

E questa è esattamente la proprietà $\mathcal{P}(n+1)$. Quindi i punti 1 e 2 del principio d'induzione sono stati verificati e possiamo concluderne che la $\mathcal{P}(n)$ vale $\forall\, n \in \mathbb{N}$.

1.4 Funzioni

Consideriamo adesso un altro importante concetto della teoria degli insiemi: le funzioni da un insieme a un altro.

Definizione 1.4.1. Dati due insiemi A e B, si dice *funzione da A a B*, una legge che a ogni elemento dell'insieme A associa uno e un solo elemento dell'insieme B. In genere le funzioni si indicano con una lettera minuscola; per indicare che f è una funzione da A a B si scriverà:

$$f : A \to B.$$

Inoltre, se $a \in A$, scriveremo $f(a) = b$, che si legge "f di a è uguale a b", per indicare che la funzione f associa all'elemento $a \in A$, l'elemento $b \in B$.

Le immagini della Figura 1.1 illustrano due esempi di funzioni:

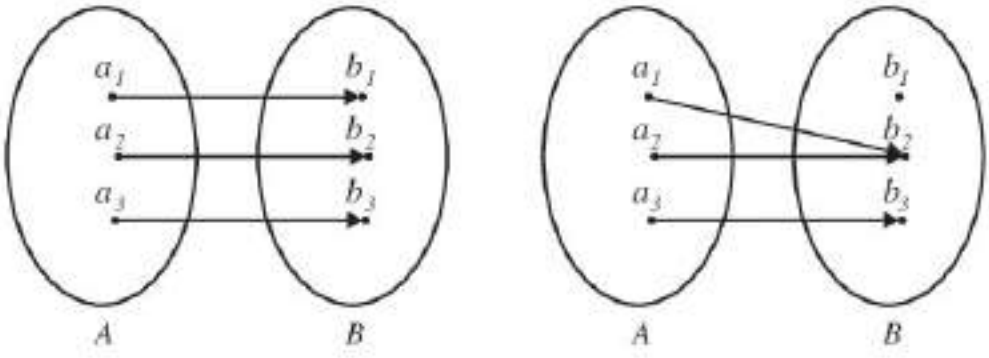


Figura 1.1: In entrambi gli esempi, gli insiemi A e B sono legati da una funzione (non la medesima).

Nel primo caso gli insiemi $A = \{a_1, a_2, a_3\}$ e $B = \{b_1, b_2, b_3\}$ hanno tre elementi ciascuno e la funzione f porta ogni elemento a_i di A nell'elemento b_i di B (per $i = 1, 2, 3$). Nel secondo caso gli insiemi sono gli stessi, ma la funzione porta a_1 e a_2 nell'elemento b_2 di B e a_3 nell'elemento b_3 (nessun elemento di A è portato in b_1). Notare che nulla vieta che una funzione porti due (o più) elementi di A nello stesso elemento di B.

In questi altri due esempi di Figura 1.2 (gli insiemi A e B sono sempre gli stessi di Figura 1.1) si hanno due leggi f che non sono funzioni: la prima di sinistra non è una funzione perché non associa niente all'elemento a_3, la seconda di destra non è una funzione perché all'elemento a_1 associa due elementi diversi di B (sia b_1 che b_2).

Vediamo altri esempi.

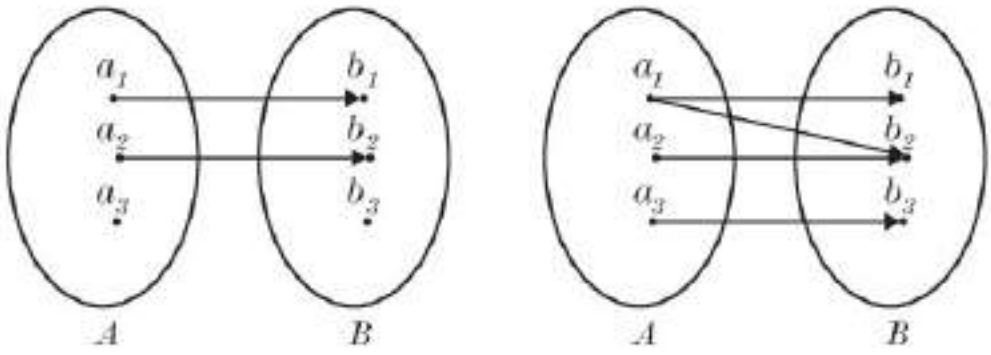


Figura 1.2: In entrambi gli esempi, la relazione che lega gli insieme A e B NON è una funzione.

Esempio 1.4.1. La legge f che a ogni numero pari (intero, positivo) associa la sua metà è una funzione dall'insieme P dei numeri pari all'insieme di tutti i numeri naturali $\mathbb{N}$;

$$f : P \to \mathbb{N}.$$

Esempio 1.4.2. Sia A l'insieme di tutte le donne (viventi e vissute) e B l'insieme di tutte le persone viventi o vissute. La legge che a ogni elemento di A associa il suo figlio o figlia (in B) è una funzione?

No, perchè non tutte le donne hanno o hanno avuto figli, quindi questa legge non assegna a ogni elemento di A un elemento di B. Inoltre alcune donne hanno o hanno avuto più figli, quindi a esse questa legge non assegna un unico elemento di B, ma più d'uno.

Se come insieme A consideriamo invece l'insieme delle madri (= donne che hanno o hanno avuto figli), allora il primo problema sarà risolto: adesso a ogni elemento di A assegniamo qualcosa in B. Questo però non basta a rendere questa relazione una funzione. Se poi inoltre invece di figli qualsiasi assegniamo a ogni madre il/la suo/a primogenito/a, allora avremo una funzione da A a B (a ogni elemento di A, una madre, corrisponde uno e un solo elemento di B, il suo primo nato).

Esempio 1.4.3. Ogni polinomio a coefficienti reali in una variabile $p(x) = a_n x^n + \cdots + a_2 x^2 + a_1 x + a_0$ rappresenta una funzione $p : \mathbb{R} \to \mathbb{R}$; si dice *grado* del polinomio la massima potenza della x che vi appare. I polinomi di grado 0 rappresentano *funzioni costanti*, che associano a ogni $x \in \mathbb{R}$ l'elemento a_0 (detto anche *termine noto* del polinomio).

Come abbiamo visto, se abbiamo $f : A \to B$, e $a \in A$, scriveremo $f(a)$ per indicare l'elemento di B ove la f "porta" l'elemento a. Ad esempio se $f : \mathbb{N} \to \mathbb{N}$ è la funzione che associa a ogni $n \in \mathbb{N}$ il numero $3n + 1$, scriveremo $f(n) = 3n + 1$ per definire la f. Inoltre per ogni

insieme A, si può definire la funzione id_A, detta *identità* su A, tale che $\forall\, a \in A$, $id_A(a) = a$.

Definizione 1.4.2. Se abbiamo due funzioni $f : A \to B$, $g : B \to C$, si può definire la funzione $g \circ f : A \to C$, *detta composizione di g e f*, definita, $\forall\, a \in A$, da:

$$(g \circ f)(a) = g(f(a)).$$

Esempio 1.4.4. Siano $f : \mathbb{R} \to \mathbb{R}$ e $g : \mathbb{R} \to \mathbb{R}$ date da $f(x) = x^3$ e $g(y) = 2y + 1$; allora si ha che $g \circ f : \mathbb{R} \to \mathbb{R}$ è data da:

$$(g \circ f)(x) = g(f(x)) = g(x^3) = 2x^3 + 1-$$

Definizione 1.4.3. Una funzione $f : A \to B$ si dice essere *iniettiva* se $\forall\, a_1, a_2 \in A$, $a_1 \neq a_2$ implica $f(a_1) \neq f(a_2)$. Si dirà invece che f è *suriettiva* se $\forall\, b \in B$, $\exists\, a \in A$ tale che $f(a) = b$. Infine si dice che f è *biunivoca* (o *bigettiva*) se è sia iniettiva che suriettiva.

Esempio 1.4.5. La funzione $f : \mathbb{R} \to \mathbb{R}$ con $f(x) = 3x + 1$ è iniettiva perché se $f(a_1) = f(a_2)$, allora $3a_1 + 1 = 3a_2 + 1$, da cui $a_1 = a_2$. La f è anche suriettiva, infatti se $b \in \mathbb{R}$, basta porre $a = [(b-1)/3]$ per avere $f(a) = b$ (verificare). Quindi la f è biunivoca.

Notare che la stessa funzione considerata da $\mathbb{N} \to \mathbb{N}$ è ancora iniettiva, ma non suriettiva.

La funzione $f : \mathbb{R} \to \mathbb{R}$ con $f(x) = x^2 + 1$ non è iniettiva, infatti ad esempio $f(3) = f(-3)$.

Dimostrare che la f non è nemmeno suriettiva.

La funzione $f : \mathbb{R}^2 \to \mathbb{R}$ con $f(x, y) = x$ è suriettiva (ad esempio $\forall\, b \in \mathbb{R}$, $f(b, 0) = b$), ma non è iniettiva (dimostrarlo per esercizio).

Osserviamo che quando una $f : A \to B$ è biunivoca, allora a ogni elemento $b \in B$ corrisponde uno e un solo $a_b \in A$, tale che $f(a_b) = b$. Si può allora anche definire la *funzione inversa* di f, e cioè una funzione $f^{-1} : B \to A$ tale che $f^{-1}(a_b) = b$.

Avremo che $f \circ f^{-1} = id_B$ e $f^{-1} \circ f = id_A$.

1.5 Operazioni, strutture algebriche

Un esempio particolare di funzioni sono le operazioni definite in un insieme. Pensiamo alla ordinaria operazione di somma in $\mathbb{N}$: non si tratta

altro che di una legge che a ogni coppia di numeri naturali ne associa un terzo, che chiamiamo somma dei primi due; si tratta cioè di una funzione:

$$\mathbb{N} \times \mathbb{N} \to \mathbb{N}.$$

Generalizziamo tale osservazione:

Definizione 1.5.1. Si dirà *operazione* (o più precisamente operazione interna) su di un insieme A una qualsiasi funzione " $*$ " che sia definita da $A \times A$ ad A. La coppia $(A, *)$ si dirà una *struttura algebrica*.

Esempio 1.5.1. Sono strutture algebriche: $(\mathbb{N}, +)$; $(\mathbb{Z}, +)$; $(\mathbb{Q}, +)$; $(\mathbb{R}, +)$; $(\mathbb{Z}, -)$; $(\mathbb{N}, \times)$; $(\mathbb{Q}, \times)$; $(\mathbb{R}, \times)$. Non sono strutture algebriche: $(\mathbb{N}, -)$; $(\mathbb{Z}, \div)$; $(\mathbb{Q}, \div)$. Infatti non si può sempre eseguire la sottrazione in $\mathbb{N}$, né la divisione in $\mathbb{Z}$; e non si può sempre dividere neanche in $\mathbb{Q}$, in quanto non si può dividere per 0! Se consideriamo l'insieme $\mathbb{Q}^* = \mathbb{Q} - \{0\}$, cioè "$\mathbb{Q}$ senza lo zero", allora $(\mathbb{Q}^*, \div)$ è una struttura.

Esercizio 1.5.1. Dire se le seguenti sono strutture algebriche: $(\mathbb{R}, \div)$, $(\mathbb{Z}^*, \div)$, $(\mathbb{R}^*, \div)$.

Definizione 1.5.2. Una struttura algebrica $(A, *)$ (ove $*$ è una qualsiasi operazione in A) si dice un *gruppo commutativo* se verifica le seguenti proprietà:

1. L'operazione è *associativa*; cioè:

$$\forall\, a, b, c \in A : a * (b * c) = (a * b) * c;$$

2. Esiste l'*elemento neutro*; cioè $\exists\, u_A \in A$, tale che:

$$u_A * a = a * u_A = a,\ \forall\, a \in A;$$

3. Esiste l'*inverso* rispetto a " $*$ ", cioè:

$$\forall\, a \in A, \exists\, a' \in A \text{ con } a' * a = a * a' = u_A;$$

4. *Commutatività*: $\forall\, a, b \in A$:

$$a * b = b * a.$$

Ad esempio $(\mathbb{Z}, +)$ è un gruppo, mentre $(\mathbb{N}, +)$ non lo è (manca l'inverso rispetto alla somma).

Sono gruppi commutativi: $(\mathbb{Z}, +)$; $(\mathbb{Q}, +)$; $(\mathbb{R}, +)$.

Se consideriamo $(\mathbb{R}, \times)$, esso non è un gruppo commutativo... perché?

Risposta: Perché lo zero non ha inverso rispetto al prodotto! Basta invece considerare $\mathbb{R}^* = \mathbb{R} - \{0\}$ e avremo che $(\mathbb{R}^*, \times)$ è un gruppo commutativo.

Definizione 1.5.3. Una struttura $(A, \oplus, \otimes)$ si dice un *anello* se valgono:

1. $(A, \oplus)$ è un gruppo commutativo;
2. In A esiste un elemento neutro rispetto a $\otimes$, cioè un elemento u tale che, per ogni $a \in A$ si abbia: $a \otimes u = u \otimes a = a$;
3. L'operazione $\otimes$ è associativa, cioè per ogni $a, b, c \in A$ si ha: $a \otimes (b \otimes c) = (a \otimes b) \otimes c$;
4. L'operazione $\otimes$ è distributiva rispetto a $\oplus$, cioè per ogni $a, b, c \in A$ si ha: $a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c)$.

Se inoltre vale anche:

5. la $\otimes$ è commutativa, cioè per ogni $a, b \in A$ si ha: $a \otimes b = b \otimes a$,

$(A, \oplus, \otimes)$ si dice un *anello commutativo*. Se poi vale anche:

6. esiste l'inverso rispetto a $\otimes$, cioè per ogni $a \in A, a \neq 0_A$, esiste $a^{-1} \in A$ tale che $a \otimes a^{-1} = u$;

allora $(A, \oplus, \otimes)$ si dice un *campo*.

Ad esempio $(\mathbb{Z}, +, \times)$ è un anello commutativo (ma non è un campo). Sono anelli commutativi: $(\mathbb{Q}, +, \times)$; $(\mathbb{R}, +, \times)$, che sono anche campi.

In particolare il campo $(\mathbb{R}, +, \times)$ sarà quello che più considereremo in questo testo; esso e altre strutture algebriche a esso collegate (come gli *spazi vettoriali*) saranno i nostri principali oggetti di studio.

Osservazione 1.5.1. Esistono anche operazioni su un insieme che coinvolgono elementi esterni all'insieme stesso; ad esempio sia $\mathbb{R}[x]$ l'insieme dei polinomi nella variabile x, a coefficienti in $\mathbb{R}$. Possiamo considerare la moltiplicazione tra elementi di $\mathbb{R}$ (numeri reali) ed elementi di $\mathbb{R}[x]$ (polinomi): questa è un'operazione su $\mathbb{R}[x]$ (perché il risultato resta in $\mathbb{R}[x]$), ma si moltiplica un elemento di $\mathbb{R}[x]$ per un elemento che non appartiene a $\mathbb{R}[x]$, un numero reale). Questa operazione è quindi una funzione $\mathbb{R} \times \mathbb{R}[x] \to \mathbb{R}[x]$. Spesso considereremo anche operazioni di questo tipo.

Osservazione 1.5.2. Infine si considerano anche operazioni su un insieme il cui risultato appartiene a un altro insieme (operazioni *esterne*),

qualcosa del tipo $A \times A \to B$. Un esempio di operazione di questo tipo è il prodotto scalare fra vettori: $\vec{\mathbf{v}} \times \vec{\mathbf{w}} = |\vec{\mathbf{v}}||\vec{\mathbf{w}}| \cos \alpha$, ove l'operazione fra vettori dà come risultato un numero reale.

1.6 Esercizi

Esercizio 1.6.1. Siano A e B i seguenti insiemi:

$$A = \{x \in \mathbb{Z} |\ x^2 \leq 12\}; \quad B = \{x \in \mathbb{N} |\ x + 1 \text{ sia primo}\};$$

Determinare $A \cap B$ e $A - B$.

Esercizio 1.6.2. Considerare i seguenti insiemi:

$$A = \{2, 5, 7, 8, 12\}; \quad B = \{1, 2, 3, 5\}; \quad C = \{7, 8\}.$$

Determinare $A \cap B$, $B \cup C$, $A - B$, $A - C$, $C - B$.

Esercizio 1.6.3. Considerare i seguenti insiemi:

$$A = \{2, 5, 7, 8\}; \quad B = \{2, 5, \{7, 8\}\}; \quad C = \{7, 8\}; \quad D = \{5\}.$$

Determinare quali delle seguenti relazioni siano vere e quali false:

$$A \subset B;\ C \subset A;\ C \subset B;\ C \in A;\ C \in B;\ D \in A;\ D \subset B;\ 5 \in D.$$

Esercizio 1.6.4. Considerare i seguenti insiemi:
$V = \{\text{Tutti i veicoli}\}$; $A = \{x \in V \mid x \text{ ha propulsione elettrica}\}$;
$B = \{x \in V |\ x \text{ sia un'auto}\}$ $C = \{x \in V \mid x \text{ ha un motore a benzina}\}$.
1) Determinare $A \cap B$, $A - B$, $A \cap C$, $B - (C \cup A)$.
2) Sia p un autobus urbano a metano e q un locomotore di un treno "Freccia Rossa" di Trenitalia. È vero che $q \in A \cup B \cup C$? E $p \in A \cup B$?

Esercizio 1.6.5. Considerare i seguenti insiemi:

$$A = \{1, 2, 5\}; \quad B = \{+, *\}.$$

Determinare $A \times B$, $B \times A$, $B^2 = B \times B$.

Esercizio 1.6.6. Sia $f : \mathbb{N} \to \mathbb{N}$ la funzione $f(n) = 2n + 1$ e dimostrare che è iniettiva ma non suriettiva.

Esercizio 1.6.7. Dire se le seguenti leggi rappresentano funzioni $\mathbb{N} \to \mathbb{N}$:

$$f(n) = 3n + 2; \quad f(n) = 2n - 1; \quad f(n) = \frac{2n+1}{3}; \quad f(n) = \frac{n(n+1)}{2}.$$

Per le f che siano funzioni dire se sono iniettive e/o suriettive.

Esercizio 1.6.8. Dire se le leggi considerate nell'esercizio precedente rappresentano funzioni quando considerate come $f : \mathbb{N} \to \mathbb{Z}$ o $f : \mathbb{N} \to \mathbb{Q}$.

Per le f che siano funzioni dire se sono iniettive e/o suriettive.

Esercizio 1.6.9. Considerare la funzione $f(x) = 3x - 1$, $f : \mathbb{R} \to \mathbb{R}$; dimostrare che è biunivoca e determinare la funzione f^{-1}.

Esercizio 1.6.10. Considerare l'operazione di sottrazione su $\mathbb{Z}$ e dire quali delle proprietà nella definizione di gruppo commutativo sono soddisfatte dalla struttura $(\mathbb{Z}, -)$.

Esercizio 1.6.11. Considerate l'operazione $* : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$, dove:

$$\forall x, y \in \mathbb{Q}, \quad x * y = xy + x + y.$$

Dire se la struttura $(\mathbb{Q}, *)$ sia un gruppo commutativo.

Esercizio 1.6.12. Dimostrare, usando il principio di induzione, che per ogni $n \in \mathbb{N}$, la somma dei numeri dispari da 1 a $2n+1$ è pari a $(n+1)^2$; cioè dimostrare che $\forall n \in \mathbb{N}$ vale la:

$$\mathcal{P}(n): \quad 1 + 3 + 5 + \cdots + (2n+1) = (n+1)^2.$$

Esercizio 1.6.13. Nell'insieme $\mathbb{R}[x]$ dei polinomi a coefficienti reali in una variabile x, consideriamo le consuete operazioni di somma e prodotto di polinomi. $(\mathbb{R}[x], +)$ è un gruppo commutativo? La struttura $(\mathbb{R}[x], +, \times)$ è un anello?

Esercizio 1.6.14. * Dire se $(A, *)$ sia una struttura algebrica nei seguenti casi:

- $A = \mathbb{Z}$, $\forall a, b \in \mathbb{Z}$: $a * b = a + b - 3$;
- $A = \mathbb{Q}$, $\forall a, b \in \mathbb{Q}$: $a * b = \frac{a}{b-1}$;
- $A = \mathbb{Q}$, $\forall a, b \in \mathbb{Q}$: $a * b = -\frac{ab}{3}$.

CAPITOLO 2

Spazi vettoriali

Prerequisiti: Strutture algebriche.

Obiettivi: Analizzare la nozione di "vettore" e darne una generalizzazione attraverso l'individuazione delle proprietà essenziali delle operazioni tra vettori e fra vettori e scalari, arrivando a definire la struttura di *spazio vettoriale* che sarà la nozione fondamentale per gli argomenti di questo corso. Impadronirsi dei concetti essenziali per lavorare con gli spazi vettoriali: basi, dimensione, sottospazi.

2.1 Introduzione: vettori geometrici

Vogliamo arrivare a introdurre un tipo di struttura algebrica, quella di spazio vettoriale, il cui scopo è di generalizzare il concetto di *vettore* come appare in geometria (o in fisica).

Consideriamo dapprima i vettori veri e propri: siano $\mathcal{V}_O^1$, $\mathcal{V}_O^2$, $\mathcal{V}_O^3$ gli insiemi dei vettori rispettivamente della retta, del piano e dello spazio euclidei, applicati in un punto fissato O.

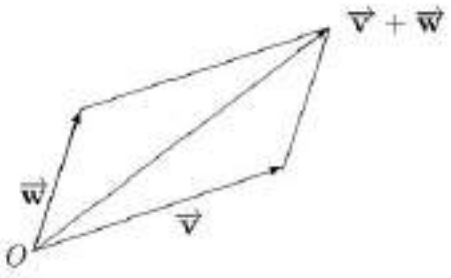
Ricordiamo che un vettore $\overrightarrow{\mathbf{v}}$ è un segmento orientato, ed è assegnato dando la sua *direzione* (la retta su cui giace), il suo *modulo* (la lunghezza del segmento, solitamente indicata con sbarre verticali: $|\overrightarrow{\mathbf{v}}|$ che è un numero reale ≥ 0) e il suo *verso* (uno dei due possibili versi di percorrenza della retta di giacitura). Osserviamo anche che se $|\overrightarrow{\mathbf{v}}| = 0$, il vettore si dice *vettore nullo* e si indica con $\mathbf{0}$. Per il vettore nullo (identificabile con il punto O) la direzione è indeterminata (esso giace su qualsiasi retta per O) e pure il verso.

In ognuno degli insiemi $\mathcal{V}_O^i$, $i = 1, 2, 3$, possiamo definire un'operazione di *somma* fra vettori:

- In $\mathcal{V}_O^1$ definiamo, $\forall\, \overrightarrow{\mathbf{v}}, \overrightarrow{\mathbf{w}} \in \mathcal{V}_O^1$, la somma $\overrightarrow{\mathbf{v}} + \overrightarrow{\mathbf{w}}$ come il vettore che ha lo stesso verso del vettore di maggior lunghezza fra $\overrightarrow{\mathbf{v}}$ e $\overrightarrow{\mathbf{w}}$

e che ha per modulo la somma $|\vec{\mathbf{v}}| + |\vec{\mathbf{w}}|$ se $\vec{\mathbf{v}}$ e $\vec{\mathbf{w}}$ sono concordi (hanno lo stesso verso) e $||\vec{\mathbf{v}}| - |\vec{\mathbf{w}}||$ se sono discordi (versi opposti).

- In $\mathcal{V}_O^2$ (e anche in $\mathcal{V}_O^3$), definiamo $\vec{\mathbf{v}} + \vec{\mathbf{w}}$ come in $\mathcal{V}_O^1$ se essi stanno su una stessa retta, altrimenti con la *regola del parallelogramma*: si considera il parallelogramma avente come lati $\vec{\mathbf{v}}$ e $\vec{\mathbf{w}}$; $\vec{\mathbf{v}} + \vec{\mathbf{w}}$ sarà dato dal vettore definito dalla diagonale del parallelogramma che esce da O (vedi figura seguente).



Si avrà che $(\mathcal{V}_O^i, +)$ risulta un *gruppo commutativo* (vedi Capitolo 1), ove l'elemento neutro è il vettore *nullo*; l'inverso di un vettore $\vec{\mathbf{v}}$ rispetto alla somma sarà il vettore $-\vec{\mathbf{v}}$, avente stessa direzione e stesso modulo di $\vec{\mathbf{v}}$, ma verso opposto. È ovvia la proprietà commutativa, mentre quella associativa si può dimostrare con una semplice costruzione geometrica (ma non lo faremo qui).

Si può definire anche un'operazione esterna fra i vettori di $\mathcal{V}_O^i$ e gli *scalari* (e cioè i numeri reali) detta *moltiplicazione per gli scalari*: $\forall\, \vec{\mathbf{v}} \in \mathcal{V}_O^i$, $\forall\, \alpha \in \mathbb{R}$, si definisce $\alpha\,\vec{\mathbf{v}}$ (si scrive anche $\alpha \cdot \vec{\mathbf{v}}$) il vettore avente:

- la stessa direzione di $\vec{\mathbf{v}}$;
- modulo pari a $|\alpha|\ |\vec{\mathbf{v}}|$;
- verso come quello di $\vec{\mathbf{v}}$ se $\alpha \geq 0$, verso opposto se $\alpha < 0$.

È facile verificare che tale prodotto (che risulta un'applicazione $\mathbb{R} \times \mathcal{V}_O^i \to \mathcal{V}_O^i$) soddisfi le seguenti proprietà :

1. $\forall\, \vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathcal{V}_O^i$, $\forall\, \alpha \in \mathbb{R}$: $\quad \alpha(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = \alpha\,\vec{\mathbf{v}} + \alpha\vec{\mathbf{w}}$.
2. $\forall\, \vec{\mathbf{v}} \in \mathcal{V}_O^i$, $\forall\, \alpha, \beta \in \mathbb{R}$: $\quad (\alpha + \beta)\vec{\mathbf{v}} = \alpha\,\vec{\mathbf{v}} + \beta\,\vec{\mathbf{v}}$.
3. $\forall\, \vec{\mathbf{v}} \in \mathcal{V}_O^i$, $\forall\, \alpha, \beta \in \mathbb{R}$: $\quad (\alpha\beta)\,\vec{\mathbf{v}} = \alpha(\beta\vec{\mathbf{v}})$.
4. $\forall\, \vec{\mathbf{v}} \in \mathcal{V}_O^i$, $\quad 1\vec{\mathbf{v}} = \vec{\mathbf{v}}$.

Vedremo che le proprietà della somma e del prodotto per gli scalari ora individuate risulteranno quelle *fondamentali* per operare con i vettori, ossia quelle che vogliamo generalizzare.

Per vedere un'altra (diversa) struttura algebrica che ha proprietà simili a quelle viste per i $\mathcal{V}_O^n$, consideriamo adesso i prodotti cartesiani di più copie dell'insieme $\mathbb{R}$ dei numeri reali, cioè gli insiemi del tipo:

$$\mathbb{R}^n = \{(a_1, a_2, \ldots, a_n) \mid \quad a_i \in \mathbb{R},\ i = 1, \ldots, n\}.$$

Su questi insiemi (convenzionalmente si pone $\mathbb{R}^1 = \mathbb{R}$) definiamo due operazioni: la *somma*, che sarà una vera e propria operazione in $\mathbb{R}^n$ e cioè una funzione $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ e l'operazione esterna di *moltiplicazione per gli scalari*, ove ancora per *scalari* si intendono gli elementi di $\mathbb{R}$ (si tratterà quindi di una funzione $\mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$).

Definiamo la somma nel modo seguente:

$\forall\, (a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$,

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$

e il prodotto per gli scalari:

$$\forall\, (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n, \forall\, \lambda \in \mathbb{R}, \qquad \lambda(a_1, a_2, \ldots, a_n) = (\lambda a_1, \lambda a_2, \ldots, \lambda a_n).$$

Con questa operazione di somma è facile vedere che $\mathbb{R}^n$ possiede una struttura di gruppo commutativo:

Proposizione 2.1.1. *La struttura algebrica $(\mathbb{R}^n, +)$ è un gruppo commutativo.*

Dimostrazione. Vediamo che $(\mathbb{R}^n, +)$ soddisfa le proprietà dei gruppi commutativi.

1. *La somma è associativa*; ciò si ricava immediatamente dal fatto che la somma è associativa in $\mathbb{R}$ e che la somma in $\mathbb{R}^n$ si esegue termine a termine.

2. *Esiste l'elemento neutro*; infatti l'elemento $\mathbf{0} = (0, 0, \ldots, 0)$, avente tutti i suoi termini nulli, verifica:

$$\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}, \qquad \forall\, \mathbf{v} = (a_1, \ldots, a_n) \in \mathbb{R}^n.$$

3. *Esiste l'inverso*: $\forall\, \mathbf{v} \in \mathbb{R}^n$, si ha
$$\mathbf{v} + (-\mathbf{v}) = \mathbf{0}, \qquad \text{ove} \qquad -\mathbf{v} = (-a_1, \ldots, -a_n).$$

4. *Commutatività*: ovviamente (dalla commutatività della somma in $\mathbb{R}$), si ha $\forall\, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$:

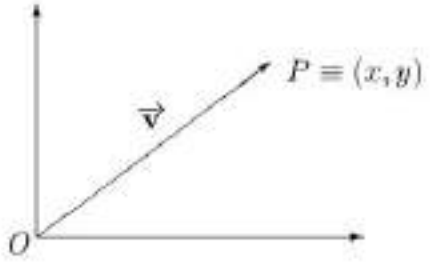
$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}.$$

□

Per quanto riguarda il prodotto per gli scalari, è facile dimostrare (lo lasciamo per esercizio) che esso verifica le proprietà analoghe a quelle del prodotto per gli scalari su $\mathcal{V}_O^i$:

1. $\forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, $\forall \alpha \in \mathbb{R}$: $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$.
2. $\forall \mathbf{v} \in \mathbb{R}^n$, $\forall \alpha, \beta \in \mathbb{R}$: $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$.
3. $\forall \mathbf{v} \in \mathbb{R}^n$, $\forall \alpha, \beta \in \mathbb{R}$: $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$.
4. $\forall \mathbf{v} \in \mathbb{R}^n$, $1\mathbf{v} = \mathbf{v}$.

Ci riferiremo spesso agli elementi di $\mathbb{R}^n$ chiamandoli *vettori*; questo ha origine nell'abituale "passaggio" che si ha in geometria analitica piana fra $\mathcal{V}_O^2$ e $\mathbb{R}^2$: possiamo associare a ogni suo elemento $(x, y) \in \mathbb{R}^2$ un vettore $\vec{\mathbf{v}} = \overrightarrow{OP}$ nel piano euclideo dotato di un sistema di coordinate cartesiane, ove O sia l'origine e P il punto di coordinate (x, y).



Viene così a stabilirsi una corrispondenza biunivoca fra l'insieme $\mathbb{R}^2$ (coppie di numeri reali), un oggetto *algebrico*, e $\mathcal{V}_O^2$, (vettori sul piano euclideo applicati nel punto O), un oggetto *geometrico*. Tale corrispondenza non è solo di insiemi: la somma in $\mathbb{R}^2$ corrisponde, alla somma che abbiamo definito in $\mathcal{V}_O^2$, e così anche la moltiplicazione per gli scalari.

Del tutto analoga è la corrispondenza tra vettori e terne di numeri reali che si ha fra $\mathbb{R}^3$ e $\mathcal{V}_O^3$ quando si studia la geometria analitica nello spazio tridimensionale.

La struttura algebrica $(\mathbb{R}^n, +, \cdot)$, dove " $\cdot$ " è la moltiplicazione per gli scalari, si indicherà d'ora in avanti semplicemente con il simbolo $\mathbb{R}^n$, che non denoterà solo l'insieme delle n-uple di reali ma l'intera struttura algebrica. Così facendo estendiamo in un certo senso il concetto di vettore a spazi diversi e, come vedremo, *n-dimensionali* (mentre i vettori "tradizionali" stanno nel piano o nello spazio della geometria euclidea). Vedremo poi in che senso $\mathbb{R}^n$ ha *dimensione* n.

Notiamo che continueremo a usare i simboli $\vec{\mathbf{v}}$, $\vec{\mathbf{w}}$ (con le "freccette") solo per i vettori veri e propri, in $\mathcal{V}_O^i$.

2.2 Spazi vettoriali

Iniziamo dando la definizione che generalizza il concetto di "vettore" a strutture algebriche formate da elementi qualsiasi ma che siano dotate di operazioni di somma e prodotto per numeri reali che soddisfino le proprietà fondamentali che abbiamo individuato per i vettori.

Definizione 2.2.1. Sia $(V,+)$ un gruppo commutativo dotato anche di un'operazione di moltiplicazione per gli elementi di $\mathbb{R}$ che soddisfi le seguenti proprietà:

1. $\forall \mathbf{v}, \mathbf{w} \in V, \forall \alpha \in \mathbb{R}$: $\alpha(\mathbf{v}+\mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$.
2. $\forall \mathbf{v} \in V, \forall \alpha, \beta \in \mathbb{R}$: $(\alpha+\beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$.
3. $\forall \mathbf{v} \in V, \forall \alpha, \beta \in \mathbb{R}$: $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$.
4. $\forall \mathbf{v} \in V$, $1\mathbf{v} = \mathbf{v}$.

Allora V si dice uno *spazio vettoriale* (su $\mathbb{R}$).

Notiamo che si potrebbe definire la struttura di spazio vettoriale non solo su $\mathbb{R}$, ma anche su un qualsiasi altro *campo* di scalari (ad esempio su $\mathbb{Q}$); noi però lavoreremo solo con spazi vettoriali su $\mathbb{R}$ e quando useremo la locuzione "spazio vettoriale" sarà sempre sottinteso che si tratta di uno spazio vettoriale su $\mathbb{R}$.

Esempio 2.2.1. Gli insiemi $\mathcal{V}_O^i$, sono ovviamente spazi vettoriali, anzi essi sono stati il nostro *prototipo* per elaborare l'idea stessa della struttura di spazio vettoriale. Analogamente, sono spazi vettoriali le strutture $\mathbb{R}^n$, $n = 1, 2, \ldots$.

Esempio 2.2.2. L'insieme $\mathbb{R}[x]$ dei polinomi nella variabile x a coefficienti in $\mathbb{R}$, è uno spazio vettoriale su $\mathbb{R}$, rispetto all'operazione di somma di polinomi e di moltiplicazione di uno scalare per un polinomio (verificare per esercizio).

Esempio 2.2.3. L'insieme $\mathcal{C}_0[\mathbb{R}]$ delle funzioni continue $\mathbb{R} \to \mathbb{R}$ è uno spazio vettoriale: la somma di due funzioni continue (ove $(f+g)(x)$ è definita da $f(x)+g(x)$) è continua, e lo stesso avviene per la funzione αf, $\forall \alpha \in \mathbb{R}$ e $\forall f \in \mathcal{C}_0[\mathbb{R}]$; verificare che $(\mathcal{C}_0[\mathbb{R}],+)$ è un gruppo è piuttosto immediato e anche che valgono le proprietà $1,\ldots,4$. (farlo per esercizio).

Esempio 2.2.4. Come per $\mathcal{C}_0[\mathbb{R}]$, sono spazi vettoriali anche $\mathcal{C}_1[\mathbb{R}]$, le funzioni $\mathbb{R} \to \mathbb{R}$ derivabili con derivata continua, e così $\mathcal{C}_2[\mathbb{R}]$, $\mathcal{C}_3[\mathbb{R}]$, $\ldots$, $\mathcal{C}_i[\mathbb{R}]$, $\ldots$ (funzioni i-volte derivabili con derivata i-esima continua) e anche $\mathcal{C}_\infty[\mathbb{R}]$ (funzioni derivabili con derivata di ogni ordine).

Esempio 2.2.5. $\mathbb{R}$ è uno spazio vettoriale su se stesso (rispetto alle usuali operazioni di somma e moltiplicazione in $\mathbb{R}$).

Esempio 2.2.6. L'insieme $\{0\} \subset \mathbb{R}$ è anch'esso uno spazio vettoriale (sempre rispetto alle usuali operazioni in $\mathbb{R}$), infatti $(\{0\}, +)$ è un gruppo additivo (verificarlo per esercizio; c'è da considerare la sola operazione possibile $0+0=0$) e la moltiplicazione per gli scalari (unica possibilità: $\alpha 0 = 0, \forall \alpha \in \mathbb{R}$) soddisfa le condizioni 1,...,4.

Esempio 2.2.7. L'insieme vuoto, $\emptyset$, NON è uno spazio vettoriale, perché non soddisfa la richiesta che esista l'elemento neutro per la somma (non avendo alcun elemento).

Esempio 2.2.8. L'insieme $\mathbb{R}^{m,n}$ delle matrici reali $m \times n$ è uno spazio vettoriale rispetto alle operazioni di somma di matrici e di moltiplicazione di una matrice per uno scalare (per le definizioni delle matrici e delle loro operazioni vedi Capitolo 3).

Proposizione 2.2.1. *In ogni spazio vettoriale si ha:* $\forall \mathbf{v} \in V$, $0 \cdot \mathbf{v} = \mathbf{0}_V$. *(Ove* $0 \in \mathbb{R}$, *mentre* $\mathbf{0}_V$ *è l'elemento neutro di* $(V, +)$*).*

Dimostrazione. Sia $\hat{\mathbf{v}}$ l'inverso in $(V, +)$ dell'elemento $0.\mathbf{v}$; avremo allora:

$$\mathbf{0}_V = \hat{\mathbf{v}} + 0 \cdot \mathbf{v} = \hat{\mathbf{v}} + (0+0) \cdot \mathbf{v} = \hat{\mathbf{v}} + 0 \cdot \mathbf{v} + 0 \cdot \mathbf{v} = (\hat{\mathbf{v}} + 0 \cdot \mathbf{v}) + 0 \cdot \mathbf{v} = \mathbf{0}_V + 0 \cdot \mathbf{v} = 0 \cdot \mathbf{v}.$$

□

Diamo ora qui di seguito alcune definizioni di cui faremo largo uso in seguito nello studio degli spazi vettoriali.

Definizione 2.2.2. Sia V uno spazio vettoriale, siano $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s \in V$. Si dice *combinazione lineare* dei $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ ogni espressione del tipo:

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_s\mathbf{v}_s, \qquad a_1, \ldots, a_s \in \mathbb{R}.$$

Le combinazioni lineari rappresentano il modo più generale di operare in uno spazio vettoriale in quanto utilizzano le due operazioni che abbiamo a disposizione in queste strutture: somme e moltiplicazioni per numeri reali.

Definizione 2.2.3. Un insieme $\mathcal{G} \subset V$ si dice un *sistema di generatori* (per lo spazio vettoriale V) se ogni $\mathbf{v} \in V$ si può scrivere come combinazione lineare di elementi di $\mathcal{G}$, cioè se comunque scelto $\mathbf{v} \in V$, esistono $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s \in \mathcal{G}$ e $a_1, \ldots, a_s \in \mathbb{R}$ tali che $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_s\mathbf{v}_s$.

La nozione di insieme di generatori è quindi importante, perché ci consente di utilizzare un ristretto numero di elementi (quelli di $\mathcal{G}$) per scrivere tutti i possibili elementi di V.

Esempio 2.2.9. Consideriamo ora, in $\mathbb{R}^2$, i seguenti vettori $\mathbf{v}_1 = (2,2)$, $\mathbf{v}_2 = (2,-2), \mathbf{v}_3 = (5,7)$. Si ha che l'insieme $\mathcal{G} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ è un insieme di generatori per $\mathbb{R}^2$, infatti ogni elemento $(a,b) \in \mathbb{R}^2$ si può scrivere come combinazione lineare di $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, ad esempio $(11,5) = \mathbf{v}_1 + 2\mathbf{v}_2 + \mathbf{v}_3$.

Osserviamo però che l'insieme $\mathcal{G} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ dell'esempio precedente non è un buon insieme di generatori per $\mathbb{R}^2$ in quanto *ridondante*: bastano infatti $\mathbf{v}_1$ e $\mathbf{v}_2$ per generare ogni elemento $(x,y) \in \mathbb{R}^2$, come si può facilmente constatare:

$$\forall\, (x,y) \in \mathbb{R}^2, \qquad (x,y) = \frac{x+y}{4}(2,2) + \frac{x-y}{4}(2,-2).$$

Ci interessa quindi un concetto più sottile di quello di insieme di generatori: vogliamo avere insiemi che non solo generino tutti i vettori di uno spazio vettoriale, ma che siano anche *efficienti* a questo rispetto, e cioè *minimali*, tali che non basti nessun loro sottoinsieme per generare ogni vettore.

2.3 Indipendenza lineare, basi

La prossima definizione svolgerà un ruolo importante nello studio degli spazi vettoriali.

Definizione 2.3.1. Dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ in uno spazio vettoriale V si dicono *linearmente indipendenti* se l'unica loro combinazione lineare che dia il vettore nullo è quella a coefficienti tutti nulli, cioè se:

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_s\mathbf{v}_s = \mathbf{0}_V \qquad \text{se e solo se} \qquad a_1 = a_2 = \cdots = a_s = 0.$$

Se i $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ non soddisfano questa proprietà allora si dicono *linearmente dipendenti*.

Un sottoinsieme $\mathcal{S} \subset V$ (anche infinito) si dice un *insieme di vettori linearmente indipendenti*, se comunque si prendano $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s \in \mathcal{S}$, essi risultano linearmente indipendenti.

Esempio 2.3.1. I vettori $\mathbf{v}_1 = (0,1)$ e $\mathbf{v}_2 = (0,3)$ sono dipendenti in $\mathbb{R}^2$, infatti:

$$3\mathbf{v}_1 - \mathbf{v}_2 = (0,0).$$

Esempio 2.3.2. I vettori $\mathbf{v_1} = (1,1)$ e $\mathbf{v_2} = (1,-1)$ sono indipendenti in $\mathbb{R}^2$, infatti se $a\mathbf{v_1} + b\mathbf{v_2} = (0,0)$, allora si ha: $(a,a) + (b,-b) = (0,0)$, cioè $(a+b, a-b) = (0,0)$:

$$\begin{cases} a+b=0 \\ a-b=0 \end{cases}$$

che ha come unica soluzione $a = b = 0$.

Esempio 2.3.3. In $\mathbb{R}[x]$ i vettori (polinomi) $\mathbf{p}(x) = 3x$ e $\mathbf{q}(x) = x^2$ sono indipendenti, infatti i soli $a, b \in \mathbb{R}$ per cui $a\mathbf{p}(x) + b\mathbf{q}(x) = bx^2 + 3ax$ sia il polinomio nullo, sono $a = b = 0$.

Osservazione 2.3.1. Se si hanno degli elementi $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ linearmente indipendenti in uno spazio vettoriale V, allora nessun $\mathbf{v}_i = \mathbf{0}_V$, altrimenti potremmo scrivere:

$$0\mathbf{v}_1 + \cdots + 1\mathbf{v}_i + \cdots + 0\mathbf{v}_s = \mathbf{0}_V,$$

e questa sarebbe una combinazione lineare nulla a coefficienti non tutti nulli.

La seguente proposizione ci mostra che la proprietà di essere linearmente indipendenti è proprio quella che ci serve per garantire, dato un insieme di generatori, che esso sia minimale (cioè che nessuno di essi sia superfluo per generare ogni vettore dello spazio).

Proposizione 2.3.1. *Siano $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s \in V$, $s \geq 2$; allora le due seguenti affermazioni sono equivalenti:*

- *$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ sono linearmente indipendenti;*
- *nessuno dei $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ è combinazione lineare dei rimanenti.*

Dimostrazione. Supponiamo che i $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ siano linearmente indipendenti; allora in effetti nessuno di loro può essere combinazione lineare degli altri, infatti se invece ci fosse un $\mathbf{v}_i$ per cui:

$$\mathbf{v}_i = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_{i-1}\mathbf{v}_{i-1} + a_{i+1}\mathbf{v}_{i+1} + \cdots + a_s\mathbf{v}_s,$$

allora

$$\mathbf{0}_V = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_{i-1}\mathbf{v}_{i-1} + (-1)\mathbf{v}_i + a_{i+1}\mathbf{v}_{i+1} + \cdots + a_s\mathbf{v}_s,$$

e i coefficienti non sono tutti nulli ($\mathbf{v}_i$ ha coefficiente -1), il che è assurdo perché i $\mathbf{v}_i$ sono indipendenti.

Viceversa, supponiamo che nessuno dei $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ sia combinazione lineare dei rimanenti; allora i vettori devono essere indipendenti, infatti se non lo fossero esisterebbe una combinazione lineare nulla:

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_s\mathbf{v}_s = \mathbf{0}_V,$$

ove almeno un coefficiente è non nullo, ad esempio, sia $a_i \neq 0$; allora potremmo ricavare (portando $a_i\mathbf{v}_i$ a destra nell'equazione e dividendo per a_i):

$$\frac{1}{a_i}(-a_1\mathbf{v}_1 - a_2\mathbf{v}_2 - \cdots - a_{i-1}\mathbf{v}_{i-1} - a_{i+1}\mathbf{v}_{i+1} - \cdots - a_s\mathbf{v}_s) = \mathbf{v}_i,$$

che è assurdo per la nostra ipotesi. □

Adesso possiamo dare la definizione che individua i "migliori" insiemi di generatori che si possano usare in uno spazio vettoriale: quelli minimali, da cui non si può togliere alcun elemento senza perdere la capacità di generare tutti i vettori dello spazio.

Definizione 2.3.2. Sia $\mathcal{B}$ un insieme di generatori di uno spazio vettoriale V. Se $\mathcal{B}$ è anche un insieme di vettori linearmente indipendenti, allora $\mathcal{B}$ si dice una *base* di V.

È di particolare importanza avere una base per uno spazio vettoriale, in quanto essa ci garantisce di poter usare un insieme minimale di vettori per poter generare tutti gli altri.

Osservazione 2.3.2. Notiamo che se V ha una base allora ne ha infinite; ad esempio se $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ è una base, allora lo sono anche $\{2\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $\{\mathbf{v}_1, \ldots, 4\mathbf{v}_n\}$, e in generale tutti gli insiemi $\{a_1\mathbf{v}_1, \ldots, a_n\mathbf{v}_n\}$, purché gli $a_1, \ldots, a_n \in \mathbb{R}$, siano tutti non nulli.

Esempio 2.3.4. I vettori $(1,0)$ e $(0,1)$ sono una base per lo spazio vettoriale $\mathbb{R}^2$, infatti sono un insieme di generatori poiché ogni $(x,y) \in \mathbb{R}^2$ si può scrivere come $x(1,0)+y(0,1)$. Sono inoltre linearmente indipendenti: se si ha: $a(1,0) + b(0,1) = (0,0)$, ciò implica che $(a,b) = (0,0)$.

L'esempio precedente si può generalizzare:

Definizione 2.3.3. Nello spazio $\mathbb{R}^n$ i vettori

$$\mathbf{e}_1 = (1,0,0,\ldots,0);\ \mathbf{e}_2 = (0,1,0,\ldots,0);\ \ldots\ ;\ \mathbf{e}_n = (0,\ldots,0,1)$$

formano una base, detta *base canonica* di $\mathbb{R}^n$ e indicata con $\mathcal{E}_n$ (o solo $\mathcal{E}$ se n si può sottintendere.)

Esercizio 2.3.1. Dimostrare che $\mathcal{E}_n = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ è una base di $\mathbb{R}^n$ (suggerimento: far vedere che $\mathbf{e}_1, \ldots, \mathbf{e}_n$ generano ogni n-upla del tipo $(a_1, \ldots, a_n) \in \mathbb{R}^n$ e che sono linearmente indipendenti, ricalcando quanto visto per $\mathbb{R}^2$).

La base canonica $\mathcal{E}_n$ è importante perché estremamente facile da usare e quindi, ove possibile, sarà sempre quella prescelta quando necessitiamo di una base per $\mathbb{R}^n$.

Analizziamo adesso il modo in cui, utilizzando una base, si possono esprimere gli elementi di uno spazio vettoriale.

Osserviamo che, fissata una base $\mathcal{B}$, il modo di generare un vettore $\mathbf{v} \in V$ tramite elementi di $\mathcal{B}$ è unico.

Proposizione 2.3.2. *Sia V uno spazio vettoriale e $\mathcal{B}$ una sua base. Allora per ogni $\mathbf{v} \in V$, l'espressione di $\mathbf{v}$ come combinazione lineare di elementi di $\mathcal{B}$ è unica, cioè sono unici gli $a_1, \dots a_n \in \mathbb{R}$ e i $\mathbf{v_1}, \dots, \mathbf{v}_n \in \mathcal{B}$ tali che: $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{v}$.*

Dimostrazione. Se la combinazione lineare non fosse unica, dovrebbero esistere $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathcal{B}$, $a_1, \dots, a_n \in \mathbb{R}$, e $b_1, \dots, b_n \in \mathbb{R}$, tali che :

$$\mathbf{v} \;=\; a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \;=\; b_1\mathbf{v_1} + \cdots + b_n\mathbf{v}_n,$$

da cui: $\mathbf{0}_V = (a_1 - b_1)\mathbf{v}_1 + \cdots + (a_n - b_n)\mathbf{v}_n$; ma i vettori $\mathbf{v}_i$ sono linearmente indipendenti in quanto elementi di una base, quindi ciò è possibile solo se si ha $a_i = b_i, \forall\, i = \mathbf{1}, \dots, n$, perciò la rappresentazione di $\mathbf{v}$ è unica. □

Osservazione 2.3.3. La proposizione precedente ci dice che ogniqualvolta si abbia uno spazio vettoriale V dotato di una base $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, a ogni elemento $\mathbf{v} \in V$ resta associata una unica n-upla di numeri reali $a_1, \dots, a_n$ tale che $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$; $a_1, \dots, a_n$ si dicono le *coordinate* di $\mathbf{v}$ rispetto a $\mathcal{B}$. Si ha così un analogo delle coordinate cartesiane in questi spazi vettoriali, e potremo lavorare in V come in $\mathbb{R}^n$.

Talvolta, se $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, per indicare $\mathbf{v}$ espresso in tali coordinate scriveremo:

$$x_{\mathcal{B}}(\mathbf{v}), \quad \text{oppure} \quad \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_{\mathcal{B}}.$$

È importante notare che quando si scrive una base, ad esempio $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, l'ordine dei vettori è molto importante in quanto ogni combinazione lineare di elementi di $\mathcal{B}$ rispetterà l'ordine dato. Ovviamente $\{\mathbf{v}_2, \mathbf{v}_1, \dots, \mathbf{v}_n\}$ non è una base diversa dalla precedente (è lo stesso insieme di vettori), ma se la scriviamo in quest'ultimo modo anziché come sopra, intendiamo che nello scrivere il vettore $x_{\mathcal{B}}(\mathbf{v})$ delle coordinate di un vettore $\mathbf{v}$ rispetto a $\mathcal{B}$, dovremo invertire l'ordine delle prime due coordinate.

Esempio 2.3.5. In $\mathbb{R}^2$, sia $\mathcal{B}_1 = \{(1,1),(1,0)\}$, che $\mathcal{B}_2 = \{(1,-1),(0,1)\}$ sono basi. Considerando l'elemento $\mathbf{v} = (3,2)$, avremo:

$$x_{\mathcal{B}_1}(\mathbf{v}) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}_{\mathcal{B}_1}; \qquad x_{\mathcal{B}_2}(\mathbf{v}) = \begin{pmatrix} 3 \\ 5 \end{pmatrix}_{\mathcal{B}_2}.$$

Mentre ovviamente, rispetto alla base canonica $\mathcal{E}_2 = \{(1,0),(0,1)\}$ avremo:

$$x_{\mathcal{B}_1}(\mathbf{v}) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}_{\mathcal{E}_2}.$$

Una volta stabilita l'importanza di avere una base per uno spazio vettoriale, possiamo porci varie domande sulle basi; dato un qualsiasi spazio vettoriale, esisterà una base per esso? Il numero di elementi di basi diverse può essere diverso? Se ho dei vettori linearmente indipendenti, posso usarli per fare una base dello spazio? I risultati seguenti ci daranno le risposte a tali domande; daremo le loro dimostrazioni solo nel caso in cui lo spazio vettoriale considerato possegga un insieme finito di generatori. Anche se tutti i risultati valgono anche per spazi con solo insiemi infiniti di generatori, in quel caso la dimostrazione è più complessa e non la affronteremo qui.

Proposizione 2.3.3. *Ogni spazio vettoriale $V \neq \{\mathbf{0}_V\}$ ammette basi.*

Dimostrazione. Come annunciato prima, dimostreremo la proposizione solo nel caso in cui V sia dotato di un insieme di generatori finito $\mathcal{G} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ (anche nel caso in cui non ci siano insiemi finiti di generatori la proposizione resta valida).

Costruiamo un sottoinsieme $\mathcal{B} \subset \mathcal{G}$ fatto di vettori indipendenti che sia massimale in $\mathcal{G}$, cioè tale che aggiungendogli un qualsiasi altro elemento $\mathbf{w}_i \in \mathcal{G} - \mathcal{B}$ si ottenga un insieme di vettori linearmente dipendenti; basterà partire dal primo $\mathbf{w}_i \neq \mathbf{0}_V$ e aggiungere il successivo se è da esso indipendente, altrimenti passare a quello dopo, e così via. Si otterrà così un insieme $\mathcal{B}$ di vettori indipendenti (che potremo scrivere $\{\mathbf{w}_1, \ldots, \mathbf{w}_s\}$, $s \leq n$, eventualmente riordinando i vettori) tale che ognuno dei restanti vettori $\mathbf{w}_{s+1}, \ldots, \mathbf{w}_n$ sia linearmente dipendente dai $\mathbf{w}_1, \ldots, \mathbf{w}_s$. Ciò implica che ognuno dei $\mathbf{w}_{s+1}, \ldots, \mathbf{w}_n$ (per la Proposizione 2.3) è generato da $\mathbf{w}_1, \ldots, \mathbf{w}_s$, quindi i vettori in $\mathcal{B}$ generano tutti quelli in $\mathcal{G}$, ma poiché $\mathcal{G}$ è un insieme di generatori per V, anche $\mathcal{B}$ lo è; inoltre $\mathcal{B}$ è un insieme di vettori linearmente indipendenti e quindi è una base.

□

Il seguente risultato ci sarà di utilità per quelli che seguono.

Proposizione 2.3.4. *Dati, in uno spazio vettoriale V, un insieme di generatori $\mathcal{G} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ e un insieme di vettori linearmente indipendenti $\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$, allora $m \geq s$.*

Dimostrazione. Poiché $\mathcal{G}$ è un insieme di generatori, abbiamo che ogni $\mathbf{v}_i$, $i = 1, \dots, s$, si può scrivere come combinazione lineare dei $\mathbf{w}_j$, $j = 1, \dots, m$:

$$\mathbf{v}_i = a_{1i}\mathbf{w}_1 + a_{2i}\mathbf{w}_2 + \cdots + a_{mi}\mathbf{w}_m.$$

Allora ogni combinazione lineare $b_1\mathbf{v}_1 + \cdots + b_s\mathbf{v}_s$ si potrà riscrivere:

$$b_1\mathbf{v}_1 + \cdots + b_s\mathbf{v}_s =$$

$$= b_1(a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m) + \cdots + b_s(a_{1s}\mathbf{w}_1 + a_{2s}\mathbf{w}_2 + \cdots + a_{ms}\mathbf{w}_m) =$$

$$= (b_1a_{11} + \cdots + b_sa_{1s})\mathbf{w}_1 + (b_1a_{21} + \cdots + b_sa_{2s})\mathbf{w}_2 + \cdots + (b_1a_{m1} + \cdots + b_sa_{ms})\mathbf{w}_m.$$

In particolare anche la combinazione lineare nulla ha tale forma. Se avessimo che ognuno dei coefficienti $b_1a_{j1} + \cdots + b_sa_{js}$ dei $\mathbf{w}_j$ fosse $= 0$, si avrebbe ovviamente che $b_1\mathbf{v}_1 + \cdots + b_s\mathbf{v}_s = \mathbf{0}_V$. Supponiamo per assurdo che valga $m < s$ e consideriamo il sistema (nelle incognite $b_1, \dots, b_s$):

$$\begin{cases} b_1a_{11} + \cdots + b_sa_{1s} = 0 \\ b_1a_{21} + \cdots + b_sa_{2s} = 0 \\ \qquad\qquad \vdots \\ b_1a_{m1} + \cdots + b_sa_{ms} = 0 \end{cases}.$$

Se $m < s$, il sistema ha più incognite che equazioni, e ciò implica che esso ha infinite soluzioni. Intuitivamente: da ogni equazione si può ricavare una variabile in funzione delle altre; alla fine ne rimarranno $s - m$ libere di variare. In modo più rigoroso, la cosa verrà vista più avanti (vedi la Proposizione 4.4.1, Teorema di Rouché-Capelli per sistemi lineari omogenei).

Si avrebbe così che il sistema ammette soluzioni $b_1, \dots, b_s$ non nulle (avendone infinite non ha solo la soluzione nulla); ma allora se $\overline{b_1}, \dots, \overline{b_s}$ è una di esse, avremmo che

$$\overline{b_1}\mathbf{v}_1 + \cdots + \overline{b_s}\mathbf{v}_s = (\overline{b_1}a_{11} + \cdots + \overline{b_s}a_{1s})\mathbf{w}_1 + \cdots + (\overline{b_1}a_{m1} + \cdots + \overline{b_s}a_{ms})\mathbf{w}_s =$$

$$= 0\mathbf{w}_1 + \cdots + 0\mathbf{w}_m = \mathbf{0}_V,$$

e questo è impossibile perchè i $\mathbf{v}_i$ sono linearmente indipendenti e i $\overline{b_1}, \dots, \overline{b_s}$ non sono tutti nulli. Quindi $m \geq s$. □

Proposizione 2.3.5. *Se $\mathcal{S} = \{\mathbf{v}_1, \dots, \mathbf{v}_s\}$ è un insieme di vettori linearmente indipendenti in uno spazio vettoriale V, è sempre possibile completare $\mathcal{S}$ a una base di V.*

Dimostrazione. Anche qui dimostreremo la Proposizione solo nel caso in cui V sia dotato di una base finita (nonostante essa valga anche nel caso in cui non ci siano basi finite di V).

Vogliamo vedere che se data una base $\mathcal{B} = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ di V (ove necessariamente $n \geq s$, dalla proposizione precedente), allora esisteranno $n - s$ vettori in $\mathcal{B}$, che, a meno di riordinamenti, si potrà indicare con $\mathbf{w}_{s+1}, \dots, \mathbf{w}_n$, tali che $\mathcal{B}' = \{\mathbf{v}_1, \dots, \mathbf{v}_s, \mathbf{w}_{s+1}, \dots, \mathbf{w}_n\}$ sia una base per V.

La dimostrazione si fa per induzione su s.

Sia dapprima $s = 1$. In questo caso il fatto che $\mathcal{S} = \{\mathbf{v}_1, \dots, \mathbf{v}_s\}$ sia un insieme di vettori linearmente indipendenti significa solo che $\mathbf{v}_1 \neq \mathbf{0}_V$. Poiché $\mathcal{B}$ è una base, avremo che esistono $a_1, \dots, a_n \in \mathbb{R}$ tali che:

$$\mathbf{v}_1 = a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \cdots + a_n\mathbf{w}_n,$$

in quanto gli a_i non possono essere tutti nulli (poiché $\mathbf{v}_1 \neq \mathbf{0}_V$). Supponiamo (a meno di un riordinamento) che $a_1 \neq 0$, quindi abbiamo che

$$\mathbf{w}_1 = \frac{1}{a_1}(\mathbf{v}_1 - a_2\mathbf{w}_2 - \cdots - a_n\mathbf{w}_n),$$

perciò l'insieme $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ è il completamento cercato. Infatti $\mathcal{B}'$ è una base poiché genera $\mathbf{w}_1$, e quindi tutto V (essa genera tutti i vettori di $\mathcal{B}$ che è una base); inoltre $\mathcal{B}'$ è un insieme di vettori indipendenti, perché se si avesse una combinazione:

$$\mathbf{0}_V = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{w}_2 + \ldots + \alpha_n\mathbf{w}_n,$$

in cui necessariamente $\alpha_1 \neq 0$, poiché $\mathbf{w}_2, \ldots, \mathbf{w}_n$ sono indipendenti, da essa si ricaverebbe :

$$\mathbf{0}_V = \alpha_1(a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \cdots + a_n\mathbf{w}_n) + \alpha_2\mathbf{w}_2 + \cdots + \alpha_n\mathbf{w}_n =$$

$$= \alpha_1 a_1\mathbf{w}_1 + (\alpha_1 a_2 + \alpha_2)\mathbf{w}_2 + \cdots + (\alpha_1 a_n + \alpha_n)\mathbf{w}_n$$

il che è impossibile perchè i $\mathbf{w}_i$ sono indipendenti e $\alpha_1 a_1 \neq 0$.

Sia ora $s > 1$. Supponiamo, per ipotesi di induzione, di sapere che l'insieme di $s-1$ vettori $\{\mathbf{v}_1, \ldots, \mathbf{v}_{s-1}\}$ può essere completato con vettori di $\mathcal{B}$ a formare una base, e cioè (a meno di riordinamenti) che $\{\mathbf{v}_1, \ldots, \mathbf{v}_{s-1}, \mathbf{w}_s, \ldots, \mathbf{w}_n\}$ sia una base. Allora potremo scrivere:

$$\mathbf{v}_s = a_1\mathbf{v}_1 + \cdots + a_{s-1}\mathbf{v}_{s-1} + a_s\mathbf{w}_s + a_{s+1}\mathbf{w}_{s+1} + \cdots + a_n\mathbf{w}_n,$$

ove almeno uno fra $a_s, \ldots, a_n$ sarà non nullo (altrimenti $\mathbf{v}_s$ sarebbe combinazione di $\mathbf{v}_1, \ldots, \mathbf{v}_{s-1}$, il che è impossibile perché $\mathcal{S}$ è un insieme di vettori indipendenti); supponiamo (a meno di riordinamenti) che sia $a_s \neq 0$. Allora si avrà :

$$\mathbf{w}_s = -\frac{1}{a_s}(a_1\mathbf{v}_1 + \cdots + a_{s-1}\mathbf{v}_{s-1} - \mathbf{v}_s + a_{s+1}\mathbf{w}_{s+1} + \cdots + a_n\mathbf{w}_n).$$

Si conclude in modo analogo a quanto fatto per $s = 1$. La base cercata sarà $\mathcal{B}' = \{\mathbf{v}_1, \ldots, \mathbf{v}_s, \mathbf{w}_{s+1}, \ldots, \mathbf{w}_n\}$, essa infatti risulta un insieme di generatori in quanto genera $\mathbf{w}_s$, e $\{\mathbf{v}_1, \ldots, \mathbf{v}_{s-1}, \mathbf{w}_s, \ldots, \mathbf{w}_n\}$ è una base. Inoltre $\mathcal{B}'$ è un insieme di vettori indipendenti in quanto se si avesse

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_{s-1}\mathbf{v}_{s-1} + \alpha_s\mathbf{v}_s + \alpha_{s+1}\mathbf{w}_{s+1} + \cdots + \alpha_n\mathbf{w}_n = \mathbf{0}_V$$

si cadrebbe in un assurdo poiché, procedendo come nel caso $s = 1$ e cioè sostituendo a $\mathbf{v}_s$ la sua espressione tramite la base $\{\mathbf{v}_1, \ldots, \mathbf{v}_{s-1}, \mathbf{w}_s, \ldots, \mathbf{w}_n\}$, si otterrebbe una combinazione lineare nulla a coefficienti non tutti nulli dei vettori di $\mathcal{S}$. □

Il prossimo risultato è importante in quanto ci dice che tutte le basi hanno lo stesso numero di elementi, e quindi che il numero delle coordinate dei vettori quando espressi come nell'Osservazione 2.3.3 non dipende dalla base scelta.

Teorema 2.3.1. *Sia V uno spazio vettoriale e $\mathcal{B}$, $\mathcal{B}'$ due basi di V. Allora $\mathcal{B}$ e $\mathcal{B}'$ possono essere messe in corrispondenza biunivoca, ossia hanno la stessa cardinalità.*

Dimostrazione. Dimostreremo anche questo teorema solo nel caso in cui le basi siano finite, anche se vale anche per basi con infiniti elementi. Osserviamo che nel caso finito l'enunciato del teorema equivale a dire che ogni base ha lo stesso numero di elementi. Consideriamo ora $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ e $\mathcal{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$. Poichè i $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ sono indipendenti e i $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ sono generatori, si ha che $m \geq n$ come immediata conseguenza della Proposizione 2.3.4. Poiché vale ovviamente anche il viceversa, avremo $n = m$. □

Il fatto che tutte le basi di uno spazio vettoriale abbiano lo stesso numero di elementi ci consente di dare la prossima definizione che individua in un certo senso la "complessità" di uno spazio vettoriale, identificando quanti elementi sono indispensabili per generare tutti gli altri.

Definizione 2.3.4. Se uno spazio vettoriale V possiede una base composta da n elementi, diremo che V ha *dimensione* n e indicheremo ciò con $\dim V = n$. Quando V non ha basi con un numero finito di elementi diremo che V ha *dimensione infinita* e scriveremo $\dim V = \infty$.

Esempio 2.3.6. Lo spazio $\mathbb{R}^n$ ha dimensione n (vedi Definizione 2.3.3).

Esempio 2.3.7. Nello spazio $\mathbb{R}[x]$, non può esistere una base finita, infatti dato un insieme finito di polinomi $\{p_1, \ldots, p_s\}$, avremo che tra loro ce ne sarà uno di grado massimo, diciamo di grado $= d$. Allora ogni combinazione lineare di $p_1, \ldots, p_s$ avrà grado $\leq d$, e non sarà possibile che $p_1, \ldots, p_s$ generino tutto $\mathbb{R}[x]$. Quindi $\dim \mathbb{R}[x] = \infty$.

Esempio 2.3.8. Riprendendo lo spazio $\mathbb{R}[x]$, abbiamo che una base (infinita) è data dai polinomi dell'insieme:

$$\mathcal{B} = \{1, x, x^2, x^3, \ldots, x^n, \ldots\}$$

infatti ogni polinomio si scrive come combinazione lineare di polinomi (monomi) in $\mathcal{B}$, e solo una combinazione a coefficienti tutti nulli di polinomi di $\mathcal{B}$ può dare il polinomio nullo.

Proposizione 2.3.6. *Se $\dim V = n$, allora n sarà il massimo numero di vettori linearmente indipendenti che sia possibile avere in V, inoltre ogni insieme di n vettori linearmente indipendenti di V sarà una base per V.*

Dimostrazione. Se si hanno $\mathbf{v}_1, \ldots, \mathbf{v}_s$ in V, linearmente indipendenti, si deve necessariamente avere $s \leq n$, per la Proposizione 2.3.4. Inoltre, se si hanno n vettori linearmente indipendenti, essi possono essere completati per ottenere una base, ma poiché il numero di elementi di una base è n stesso, necessariamente i vettori dati dovevano già formare una base. □

Esempio 2.3.9. Consideriamo, in $\mathbb{R}^3$, $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, ove si abbia $\mathbf{v}_1 = (1, 2, 3)$, $\mathbf{v}_2 = (4, 5, 6)$ e $\mathbf{v}_3 = (0, 1, 0)$. $\mathcal{B}$ è una base per $\mathbb{R}^3$?

Per la proposizione precedente, poiché $\dim \mathbb{R}^3 = 3$, basta controllare se i tre vettori siano linearmente indipendenti. Da

$$a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3 = (0, 0, 0)$$

si ricava:

$$\begin{cases} a + 4b = 0 \\ 2a + 5b + c = 0 \\ 3a + 6b = 0 \end{cases}$$

da cui si ottiene $a = b = c = 0$, e quindi i tre vettori sono indipendenti e $\mathcal{B}$ è una base.

2.4 Sottospazi

Consideriamo adesso le "sottostrutture" di uno spazio vettoriale, cioè i suoi sottospazi:

Definizione 2.4.1. Sia V uno spazio vettoriale; un suo sottoinsieme $W \subset V$ si dice *sottospazio vettoriale* di V se W è uno spazio vettoriale rispetto alle operazioni (somma e prodotto per gli scalari) di V.

Proposizione 2.4.1. *Un sottoinsieme W di uno spazio vettoriale V è un sottospazio vettoriale se e solo se valgono le seguenti proprietà:*

1. $\mathbf{0}_V \in W$;
2. $\forall\, \mathbf{w}_1, \mathbf{w}_2 \in W, \forall\, \alpha, \beta \in \mathbb{R}, \qquad \alpha\mathbf{w}_1 + \beta\mathbf{w}_2 \in W$.

Dimostrazione. Se W è un sottospazio, le due proprietà valgono ovviamente. Consideriamo l'inverso; vogliamo vedere che se valgono la 1 e la 2 allora W deve essere un sottospazio vettoriale.

Osserviamo innanzitutto che la 2 implica che sommare elementi di W e moltiplicare elementi di W per scalari fa rimanere in W, e cioè che queste risultano operazioni anche su W.

Per vedere che $(W, +)$ è un gruppo commutativo basta osservare che le proprietà associativa e commutativa valgono in W perché valgono in tutto V; che esiste l'elemento neutro in W per la 1 e che $\forall\, \mathbf{w} \in W$ esiste l'inverso in W poiché $(-1)\mathbf{w} \in W$ per la 2 e

$$(-1)\mathbf{w} + \mathbf{w} = (-1 + 1)\mathbf{w} = 0\mathbf{w} = \mathbf{0}_V.$$

Concludiamo osservando che le proprietà del prodotto per gli scalari valgono in W in quanto valgono in tutto V. □

Esempio 2.4.1. In $\mathbb{R}^2$, sia $W = \{(x, y) | y = 0\}$; W è un sottospazio di $\mathbb{R}^2$? Sicuramente $(0, 0) \in W$, poi se ho due elementi $(a, 0), (b, 0) \in W$, avremo: $\alpha(a, 0) + \beta(b, 0) = (\alpha a + \beta b, 0) \in W$, e quindi W è un sottospazio.

Esempio 2.4.2. In ogni spazio vettoriale V, V stesso e $\{\mathbf{0}_V\}$ sono suoi sottospazi vettoriali (si dicono *sottospazi banali* di V).

Esempio 2.4.3. L'insieme $W = \{(x, y) |\ y^3 = x\}$ non è un sottospazio di $\mathbb{R}^2$; anche se la 1. è soddisfatta, la 2. non vale; prendiamo ad esempio $(8, 2), (1, 1) \in W$, si ha: $(8, 2) + (1, 1) = (9, 3)$ che non sta in W.

Esempio 2.4.4. L'insieme $W = \{(x, y, z) |\ 3 = x - y + z\}$ non è un sottospazio di $\mathbb{R}^3$, infatti $(0, 0, 0)$ non sta in W.

Esempio 2.4.5. L'insieme $W = \{p(x) \mid \text{grado di } p(x) \leq 3\}$ è un sottospazio di $\mathbb{R}[x]$: infatti il polinomio nullo gli appartiene e ogni combinazione lineare di polinomi di grado ≤ 3 ha ancora grado ≤ 3.

Poiché un sottospazio W di uno spazio vettoriale V è un sottoinsieme che "eredita" da V, tramite le operazioni di V, la struttura di spazio vettoriale, si ha quindi che le nozioni di sistema di generatori, base, dimensione, si applicano naturalmente anche ai sottospazi.

In particolare se $W \subset V$ è un sottospazio e $W \neq V$, avremo sempre che $\dim W < \dim V$ (infatti una base di W si potrà sempre estendere a una base di V, per la Proposizione 2.3.5).

Definizione 2.4.2. Sia V uno spazio vettoriale; dati $\mathbf{v}_1, \ldots, \mathbf{v}_s \in V$ si definisce *sottospazio generato* da $\mathbf{v}_1, \ldots, \mathbf{v}_s$ lo spazio:

$$\langle \mathbf{v}_1, \ldots, \mathbf{v}_s \rangle = \{\mathbf{v} \in V |\ \exists\, a_1, \ldots, a_s \in \mathbb{R} \text{ tali che } \mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_s\mathbf{v}_s\}.$$

Quindi $\langle \mathbf{v}_1, \ldots, \mathbf{v}_s \rangle$ risulta lo spazio di tutte le combinazioni lineari dei $\mathbf{v}_1, \ldots, \mathbf{v}_s$ (in alcuni testi questo viene indicato con $\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_s)$).

Notiamo che in modo ovvio si ha che $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$ è un insieme di generatori per $\langle \mathbf{v}_1, \ldots, \mathbf{v}_s \rangle$; ma non è detto che sia una base, perché i $\mathbf{v}_1, \ldots, \mathbf{v}_s$ potrebbero non essere linearmente indipendenti.

Esempio 2.4.6. In $\mathbb{R}^3$, consideriamo:

$$\mathbf{v}_1 = (2, 2, 0); \qquad \mathbf{v}_2 = (1, 0, 0); \qquad \mathbf{v}_3 = (0, 2, 0).$$

Si ha:

$$\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle = \{a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3 |\ \forall\, a, b, c \in \mathbb{R}\} =$$
$$= \{(2a + b, 2a + 2c, 0) |\ \forall\, a, b, c \in \mathbb{R}\}.$$

Sebbene $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$ sia stato definito usando tre vettori, abbiamo che $\dim\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle = 2$, in quanto i vettori $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ non sono linearmente indipendenti (per esempio $\mathbf{v}_1 = 2\mathbf{v}_2 + \mathbf{v}_3$) e bastano due di essi ($\mathbf{v}_2$ e $\mathbf{v}_3$ ad esempio) per generare tutto $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$, infatti:

$$(2a + b, 2a + 2c, 0) \ = \ (2a + b)\mathbf{v}_2 \ + \ (a + c)\mathbf{v}_3.$$

Si ha quindi che $\{\mathbf{v}_2, \mathbf{v}_3\}$ risulta una base per $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$, e

$$\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle.$$

In generale, se $\mathbf{v}_1, \ldots, \mathbf{v}_s$ sono linearmente indipendenti, lo spazio $\langle \mathbf{v}_1, \ldots, \mathbf{v}_s \rangle$ ha $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$ come base, altrimenti si trova una sua base prendendo il massimo numero di vettori indipendenti fra i $\mathbf{v}_1, \ldots, \mathbf{v}_s$.

Esempio 2.4.7. Siano $\mathbf{v}_1 = (0,1,2,3)$, $\mathbf{v}_2 = (1,1,1,1)$ e $\mathbf{v}_3 = (1,2,3,0)$ in $\mathbb{R}^4$; quale dimensione ha $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$?

Vediamo che i tre vettori sono linearmente indipendenti; siano $a, b, c \in \mathbb{R}$ tali che:

$$a(0,1,2,3) + b(1,1,1,1) + c(1,2,3,0) = \mathbf{0};$$

che equivale a

$$(b+c, a+b+2c, 2a+b+3c, 3a+b) = (0,0,0,0).$$

Risolvendo il sistema (farlo per esercizio):

$$\begin{cases} b+c=0 \\ a+b+2c=0 \\ 2a+b+3c=0 \\ 3a+b=0 \end{cases},$$

si ottiene $a = b = c = 0$, quindi l'unica combinazione lineare nulla dei tre vettori risulta quella a coefficienti nulli, perciò $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ è una base di $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle$ e $\dim \langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \rangle = 3$.

Vediamo adesso alcune operazioni fra sottospazi:

Proposizione 2.4.2. *Se V_1, V_2 sono sottospazi di V, allora anche $V_1 \cap V_2$ lo è.*

Dimostrazione. Utilizzando la Proposition 2.4.1, vediamo che $\mathbf{0}_V \in V_1 \cap V_2$, in quanto $\mathbf{0}_V \in V_1$ e $\mathbf{0}_V \in V_2$. Se poi si hanno $\mathbf{w}_1, \mathbf{w}_2 \in V_1 \cap V_2$, ogni combinazione lineare $a\mathbf{w}_1 + b\mathbf{w}_2 \in V_1 \cap V_2$, in quanto $\mathbf{w}_1, \mathbf{w}_2 \in V_1$ e $\mathbf{w}_1, \mathbf{w}_2 \in V_2$. □

Osservazione 2.4.1. Osserviamo che invece, in generale anche se V_1, V_2 sono sottospazi di uno spazio vettoriale V, l'insieme $V_1 \cup V_2$, NON è uno sottospazio di V, come il seguente esempio mostra.

Esempio 2.4.8. Consideriamo, in $\mathbb{R}^2$, gli insiemi: $V_1 = \{(x,y) \in \mathbb{R}^2 \,|\, x = 0\}$, $V_2 = \{(x,y) \in \mathbb{R}^2 \,|\, y = 0\}$ sono sottospazi, ma l'insieme $V_1 \cup V_2 = \{(x,y) \in \mathbb{R}^2 \,|\, x = 0 \text{ oppure } y = 0\}$ non è un sottospazio. Infatti se consideriamo $\mathbf{w}_1 = (3,0)$, $\mathbf{w}_2 = (0,5)$ essi stanno in $V_1 \cup V_2$, ma $\mathbf{w}_1 + \mathbf{w}_2 = (3,5) \notin V_1 \cup V_2$.

Abbiamo invece un'operazione di *somma* tra sottospazi:

Definizione 2.4.3. Se $V_1, \ldots, V_s$ sono sottospazi di uno spazio vettoriale V, si dice *spazio somma* di $V_1, \ldots, V_s$ il sottospazio

$$V_1 + V_2 + \cdots + V_s = \{\mathbf{v} \in V \,|\, \exists \mathbf{v}_i \in V_i,\ i = 1, \ldots, s,\ \mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_s\}.$$

Osservazione 2.4.2. Si ha che $V_1 + V_2 + \cdots + V_s$ è il *minimo* sottospazio di V contenente $V_1, V_2, \ldots, V_s$ (ovviamente ogni $V_i \subset V_1 + V_2 + \cdots + V_s$; se invece W è un sottospazio di V che contiene ogni V_i, allora contiene anche le combinazioni lineari dei loro vettori, quindi $V_1 + V_2 + \cdots + V_s \subset W$).

Riguardo alla somma di sottospazi si ha la seguente:

Proposizione 2.4.3 (Formula di Grassmann). *Siano V_1, V_2 sottospazi dello spazio vettoriale V. Allora si ha:*

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Dimostrazione. Eseguiamo la dimostrazione nel caso in cui le dimensioni di V_1 e V_2 siano finite. Sia $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$ una base di $V_1 \cap V_2$; per la Proposizione 2.3.5, si può completare $\mathcal{B}$ sia a $\mathcal{B}_1$, base di V_1, che a $\mathcal{B}_2$, base di V_2, ove:

$$\mathcal{B}_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_s, \mathbf{w}_{s+1}, \ldots, \mathbf{w}_{s+r}\},\ \mathcal{B}_2 = \{\mathbf{v}_1, \ldots, \mathbf{v}_s, \mathbf{w}'_{s+1}, \ldots, \mathbf{w}'_{s+t}\}$$

(abbiamo posto $\dim V_1 = s + r$, $\dim V_2 = s + t$).

Ovviamente $\mathcal{B}_3 = \{\mathbf{v}_1, \ldots, \mathbf{v}_s, \mathbf{w}_{s+1}, \ldots, \mathbf{w}_{s+r}, \mathbf{w}'_{s+1}, \ldots, \mathbf{w}'_{s+t}\}$ è un insieme di generatori per $V_1 + V_2$; vediamo che ne è anche una base. Consideriamo una combinazione lineare nulla:

$$a_1\mathbf{v}_1 + \cdots + a_s\mathbf{v}_s + b_{s+1}\mathbf{w}_{s+1} + \cdots + b_{s+r}\mathbf{w}_{s+r} + c_{s+1}\mathbf{w}'_{s+1} + \cdots + c_{s+t}\mathbf{w}'_{s+t} = \mathbf{0}_V.$$

Da essa si ricava:

$$a_1\mathbf{v}_1 + \cdots + a_s\mathbf{v}_s + b_{s+1}\mathbf{w}_{s+1} + \cdots + b_{s+r}\mathbf{w}_{s+r} = -c_{s+1}\mathbf{w}'_{s+1} - \cdots - c_{s+t}\mathbf{w}'_{s+t}$$

e quindi $-c_{s+1}\mathbf{w}'_{s+1} - \cdots - c_t\mathbf{w}'_{s+t} \in V_1$, cosicché appartiene anche a $V_1 \cap V_2$. Ma ciò è possibile solo se il vettore $-c_{s+1}\mathbf{w}'_{s+1} - \cdots - c_{s+t}\mathbf{w}'_{s+t} =$

$\mathbf{0}_V$, perché altrimenti sarebbe generato da $\mathbf{v_1}, \ldots, \mathbf{v}_s$, mentre i $\mathbf{w}'_i$ sono indipendenti dai $\mathbf{v}_j$ (insieme formano una base di V_2).

Per l'indipendenza dei $\mathbf{w}'_i$ si ha che $c_{s+1} = \cdots = c_{s+t} = 0$. In modo analogo si ottiene che $b_{s+1} = \cdots = b_{s+r} = 0$, quindi la prima uguaglianza diviene $a_1\mathbf{v_1} + \cdots + a_s\mathbf{v}_s = \mathbf{0}_V$, che dà ancora $a_1 = \cdots = a_s = 0$ e quindi $\{\mathcal{B}_3\}$ è una base di $V_1 + V_2$ e $\dim(V_1 + V_2) = s+r+t = (s+r)+(s+t)-s$, che è l'uguaglianza cercata. □

Nel caso particolare in cui $V_1 \cap V_2 = \{\mathbf{0}\}_V$, avremo quindi che la somma "rispetta le dimensioni": $\dim(V_1+V_2) = \dim V_1 + \dim V_2$. Questo caso si denota nel modo seguente.

Definizione 2.4.4. Siano V_1,V_2 sottospazi di uno spazio vettoriale V; se $V_1 \cap V_2 = \{\mathbf{0}_V\}$, lo spazio somma $V_1 + V_2$ si dice *somma diretta* di V_1,V_2 e si indica con $V_1 \oplus V_2$.

Dalla Proposizione 2.4.3 si ha:

$$\dim V_1 \oplus V_2 = \dim V_1 + \dim V_2.$$

2.5 Esercizi

Esercizio 2.5.1. Dati, in $\mathbb{R}^3$, i vettori $\mathbf{v_1} = (\frac{1}{2}, 0, 1)$, $\mathbf{v}_2 = (1, 1, -1)$, $\mathbf{v}_3 = (2, 1, 0)$, sia $\mathbf{u} = 2\mathbf{v_1} + 2\mathbf{v}_2 - \frac{3}{2}\mathbf{v}_3$. Scrivere $\mathbf{u}$ come combinazione lineare dei seguenti vettori: $\mathbf{w}_1 = (0, 0, 3)$, $\mathbf{w}_2 = (0, 1, 1)$, $\mathbf{w}_3 = (-1, 1, 0) \in \mathbb{R}^3$.

Esercizio 2.5.2. Sia L una retta nello spazio tridimensionale, passante per il punto O. Verificare che l'insieme $\mathcal{V}_L$ formato dai vettori di $\mathcal{V}^3_O$ contenuti in L è uno sottospazio vettoriale di $\mathcal{V}^3_O$.

Esercizio 2.5.3. Sia V uno spazio vettoriale e $\mathbf{v} \neq \mathbf{0}$ un suo elemento. Verificare che l'insieme $W = \{\mathbf{w} \in V \mid \mathbf{w} = \alpha\mathbf{v}\}$, formato dai multipli di $\mathbf{v}$, è uno sottospazio vettoriale di V.

Esercizio 2.5.4. * Considerare le strutture $(\mathbb{R}^2, +, \times_1)$, $(\mathbb{R}^2, +, \times_2)$ ove la somma sia quella usuale, ma l'operazione di prodotto fra numeri reali ed elementi di $\mathbb{R}^2$ sia definita da:

$$\lambda \times_1 (a, b) = (\lambda^2 a,\ \lambda^2 b), \qquad \forall \lambda \in \mathbb{R},\ (a, b) \in \mathbb{R}^2;$$

oppure da:

$$\lambda \times_2 (a, b) = (2\lambda a, 2\lambda b), \qquad \forall \lambda \in \mathbb{R},\ (a, b) \in \mathbb{R}^2;$$

Le strutture $(\mathbb{R}^2, +, \times_1)$ e $(\mathbb{R}^2, +, \times_2)$ sono spazi vettoriali?

Esercizio 2.5.5. Considerare la struttura $(\mathbb{Z}, +, \times)$, ove $\mathbb{Z}$ è l'insieme dei numeri interi, la somma è quella usuale, ma l'operazione " $\times$ " di prodotto fra numeri reali ed elementi di $\mathbb{Z}$ sia definita da:

$$\alpha \times z = 0, \qquad \forall \alpha \in \mathbb{R},\ \forall z \in \mathbb{Z};$$

La struttura $(\mathbb{Z}, +, \times)$ è uno spazio vettoriale?

Esercizio 2.5.6. Dire se i seguenti sottoinsiemi sono anche sottospazi (nei rispettivi $\mathbb{R}^n$):

$$W = \{(x, y, z) \mid x - 3y + 4z = 0\};$$

$$W = \{(x, y) \mid x^2 + y^2 = 0\}, \quad W = \{(x, y, z) \mid x - y = 0, x + z = 0\};$$

$$W = \{(x, y, z, t) \mid 3x - t = y + z, x - y + z = 0\};$$

$$W = \{(x, y, z) \mid x^2 - 3y = 0\};$$

$$W = \{(x, y) \mid x^2 = y^2\};$$

$$W = \{(x, y) \mid x = \sin y\};$$

$$W = \{(x, y, z) \mid \sin^2 x + \sin^2 y - 1 = z\}.$$

Esercizio 2.5.7. Stabilire se i seguenti sottoinsiemi di $V = \mathbb{R}^3$ sono dei sottospazi di V:

$$S_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 = 0\},$$

$$S_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 = 1\},$$

$$S_3 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 \geq 0\}.$$

Esercizio 2.5.8. Verificare che il seguente insieme è un sottospazio di $\mathbb{R}^4$ e calcolarne una base:

$S = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + 3x_2 - x_3 + 4x_4 = 0\}$.

Esercizio 2.5.9. Dimostrare che l'insieme delle soluzioni del sistema lineare

$$\begin{cases} 3x - 2y - z = 0 \\ 2x - y - z = 0 \end{cases}$$

è uno sottospazio vettoriale di $\mathbb{R}^3$.

Esercizio 2.5.10. Sia V un R-spazio vettoriale di vettori applicati in P_0. Siano π un piano passante per P_0 e r, s due rette per P_0. Stabilire se i seguenti sottoinsiemi di V sono sottospazi:

- $S_\pi = \{\overrightarrow{PP_0} \in V \mid P \in \pi\}$,
- $S_r = \{\overrightarrow{PP_0} \in V \mid P \in r\}$,
- $S_s = \{\overrightarrow{PP_0} \in V \mid P \in s\}$,
- $S_r \cup S_s$.

Esercizio 2.5.11. * Dimostrare che due vettori in $\mathcal{V}_O^2$ sono linearmente indipendenti se e solo se non giacciono sulla stessa retta.

Esercizio 2.5.12. * Dimostrare che tre vettori in $\mathcal{V}_O^3$ sono linearmente indipendenti se e solo se non sono complanari.

Esercizio 2.5.13. Dire, al variare di $k \in \mathbb{R}$, se i seguenti vettori siano una base per $\mathbb{R}^3$:

$$\mathbf{v_1} = (1, -1, 0),\ \mathbf{v_2} = (1, k, 1),\ \mathbf{v_3} = (3, 1, k-2).$$

Esercizio 2.5.14. * Siano $V = \langle \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3} \rangle$ e $W = \langle \mathbf{w_1}, \mathbf{w_2}, \mathbf{w_3} \rangle$ due sottospazi di $\mathbb{R}^4$ con $\mathbf{v_1} = (1, 1, 0, -1)$, $\mathbf{v_2} = (1, 2, 3, 0)$, $\mathbf{v_3} = (2, 3, 3, -1)$, $\mathbf{w_1} = (1, 2, 2, -2)$, $\mathbf{w_2} = (2, 3, 2, -3)$, $\mathbf{w_3} = (1, 3, 4, -3)$. Determinare la dimensione di $V + W$.

Esercizio 2.5.15. Siano $V = \{\overrightarrow{v} \in \mathcal{V}_O^3 \mid |\overrightarrow{v}| \leq 1\}$ (V è dato dai vettori contenuti in una sfera di centro O e raggio $=1$). V è un sottospazio di $\mathcal{V}_O^3$? In caso affermativo, determinare la dimensione di V.

Esercizio 2.5.16. Stabilire se le due seguenti terne di vettori siano linearmente indipendenti in $\mathbb{R}^4$ e, se lo sono, completarli a una base di $\mathbb{R}^4$:

- $\mathbf{v_1} = (1, 0, 0, 0)$, $\mathbf{v_2} = (0, 1, 0, 0)$, $\mathbf{v_3} = (2, -1, 0, 0)$;
- $\mathbf{w_1} = (1, 0, 0, 0)$, $\mathbf{w_2} = (1, 1, 0, 0)$, $\mathbf{w_3} = (0, 1, 2, 1)$.

Esercizio 2.5.17. Verificare che i seguenti vettori formano un sistema di generatori di $\mathbb{R}^3$ ma non una base:
$\mathbf{v_1} = (-1, 0, 0)$, $\mathbf{v_2} = (0, 1, 1)$, $\mathbf{v_3} = (0, 0, 1)$, $\mathbf{v_4} = (1, 1, 1)$.

Esercizio 2.5.18. Trovare un sistema di generatori di $\mathbb{R}^5$ che non sia una base.

Esercizio 2.5.19. Determinare quali tra i seguenti sottoinsiemi di $\mathbb{R}^4$ sono sottospazi e trovarne eventualmente una base:

$$W_1 = \{(a, b, 2a, -b) \in \mathbb{R}^4 \mid a, b \in \mathbb{R}\},$$
$$W_2 = \{(a, 1+a, a-b, 0) \in \mathbb{R}^4 \mid a, b \in \mathbb{R}\},$$
$$W_3 = \{(a+b, b-a, 0, -b) \in \mathbb{R}^4 \mid a, b \in \mathbb{R}\},$$
$$W_4 = \{(a, 0, b, ab) \in \mathbb{R}^4 \mid a, b \in \mathbb{R}\}.$$

Esercizio 2.5.20. Si consideri il seguente sottospazio di $\mathbb{R}^3$:
$V = \{(x, y, z) \in \mathbb{R}^3 \mid x - ky + z = 0,\ kx - y + z = 0\}$.
Si determini $\dim V$ al variare del parametro reale k.

Esercizio 2.5.21. * Si considerino i seguenti sottospazi di $\mathbb{R}^4$:
$V = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - ky = z = 0\}$,
$W = \{(x, y, z, t) \in \mathbb{R}^4 \mid t + y = kx + z = 0\}$
dove k è un parametro reale. Dopo aver scritto una base per V e una base per W, determinare i valori di k per cui risulta $W + V = \mathbb{R}^4$.

Esercizio 2.5.22. * Dimostrare che il sottoinsieme

$$V = \{p(x) \in \mathbb{R}[x] \mid p(0) = 0\} \subset \mathbb{R}[x]$$

è un sottospazio di $\mathbb{R}[x]$ e che $\dim V = \infty$.

Esercizio 2.5.23. * Il sottoinsieme

$$V = \{p(x) \in \mathbb{R}[x] \mid p(x) \text{ sia divisibile per } x\} \subset \mathbb{R}[x]$$

è un sottospazio di $\mathbb{R}[x]$?

Esercizio 2.5.24. Trovare una base del sottospazio
$S = \{(x, y, z) \in \mathbb{R}^3 \mid x + y = 0\} \subset \mathbb{R}^3$.

Esercizio 2.5.25. Trovare una base di $V \cap V' \subset \mathbb{R}^3$ dove
$V = \{(x, y, z) \in \mathbb{R}^3 \mid x - y = 0\}$ e $V' = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y - z = 0\}$.

Esercizio 2.5.26. Trovare una base del sottospazio $S + S' \subset \mathbb{R}^4$ dove
$S = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y = 0\}$ e $S' = T \cap T'$, ove

$$T = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - 2z = 0\}\ ,\ T' = \{(x, y, z, t) \in \mathbb{R}^4 \mid y - z = 0\}.$$

Esercizio 2.5.27. * Sia $S = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y, z) = (z, y, x)\}$.

- Dimostrare che S è un sottospazio di $\mathbb{R}^3$;
- Determinare una base di S;
- Determinare le coordinate del vettore $(7, 2, 7)$ rispetto alla base di S scelta nel punto precedente.

Esercizio 2.5.28. Fare un esempio di un sottospazio di $\mathbb{R}^3$ che abbia dimensione 2 e che contenga il vettore $(2, -1, 1)$.

Esercizio 2.5.29. Fare un esempio di un sottospazio di $\mathbb{R}^4$ che abbia dimensione 2 e che contenga il vettore $(2, -1, 1, 0)$.

Esercizio 2.5.30. * Avvertenza: Per svolgere questo esercizio bisogna conoscere il contenuto del Capitolo 3 (matrici).

Sia $V = \mathbb{R}^{3,3}$. Siano S il sottoinsieme delle matrici simmetriche in V e S' il sottoinsieme delle matrici antisimmetriche di V (una matrice A è simmetrica se $A = {}^tA$ e antisimmetrica se $A = -({}^tA)$). Verificare che S e S' sono sottospazi di V e dimostrare che $S \oplus S' = V$.

Suggerimento. Per la seconda domanda, utilizzare la Proposizione 2.4.3.

Esercizio 2.5.31. Dare un esempio di spazio vettoriale che non sia un $\mathbb{R}^n$ e abbia dimensione ≥ 4.

Esercizio 2.5.32. Dire, al variare di $n \in \mathbb{N}$, $n \neq 0$, se $\mathbb{R}^n$ sia un sottospazio vettoriale di $\mathbb{R}^{n+2}$.

Esercizio 2.5.33. Dare un esempio di un sottospazio vettoriale $W \subset \mathbb{R}^4$ che abbia dimensione d, con $1 \leq d \leq 3$.

Esercizio 2.5.34. Dimostrare che per un sottospazio vettoriale $W \subset V$ che abbia dimensione pari alla $\dim V$, si ha necessariamente che $W = V$.

Esercizio 2.5.35. Disegnare i seguenti sottoinsiemi di $\mathbb{R}^2$ e dire se sono dei sottospazi:

1. $\{(x, y) \in \mathbb{R}^2 \mid x + y = 0\}$,
2. $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y = 0\}$,
3. $\{(x, y) \in \mathbb{R}^2 \mid x + y - 2 = 0\}$.

Esercizio 2.5.36. Mostrare con un esempio grafico in R^2 che l'unione di due sottospazi non è un sottospazio a meno che l'uno non sia contenuto nell'altro.

Esercizio 2.5.37. Mostrare con un esempio grafico in R^3 che l'unione di due sottospazi non è un sottospazio a meno che l'uno non sia contenuto nell'altro.

Esercizio 2.5.38. Mostrare con un esempio grafico in R^2 che la somma di due sottospazi è un sottospazio.

Esercizio 2.5.39. Mostrare con un esempio grafico in R^3 che la somma di due sottospazi è un sottospazio.

Esercizio 2.5.40. Disegnare i seguenti sottoinsiemi di $\mathbb{R}^3$ e dire se sono dei sottospazi:

1. $\{(x, y) \in \mathbb{R}^2 \mid x + y - 2z = 0\}$,
2. $\{(x, y) \in \mathbb{R}^2 \mid x + y - 2z = 0 = x\}$,
3. $\{(x, y) \in \mathbb{R}^2 \mid x + y + 3z^3 = 0\}$,
4. $\{(x, y) \in \mathbb{R}^2 \mid x + y - 5 = 0\}$,
5. $\{(x, y) \in \mathbb{R}^2 \mid x + y - 1 = 0 = x - y + z\}$.

CAPITOLO 3

Le matrici

Prerequisiti: Nessun prerequisito per le generalità sulle matrici e per le loro operazioni; conoscenza delle strutture algebriche (gruppo e anello, vedi Capitolo 1) per l'analisi della struttura di $(\mathbb{R}^{m,n}, +)$ e $(\mathbb{R}^{n,n}, +, \cdot)$; conoscenza degli spazi vettoriali (definizione, basi) solo per le osservazioni sullo spazio delle righe e delle colonne di una matrice (alla fine della Sezione 3.1).

Obiettivi: Acquisizione della familiarità con le operazioni fra matrici, la loro simbologia e le nozioni fondamentali.

3.1 Introduzione

In questo capitolo si introdurranno le matrici che costituiscono uno degli strumenti principali dell'algebra lineare (come si potrà verificare nei capitoli successivi). Vedremo come le matrici sono uno strumento naturale per immagazzinare e utilizzare dei dati.

Da un punto di vista più algebrico, definiremo varie operazioni negli insiemi delle matrici e considereremo quali strutture si formino e che proprietà posseggano.

3.2 Generalità sulle matrici

Con la parola *matrice* intenderemo una tabella rettangolare di numeri, come le seguenti:

$$\begin{pmatrix} 1 & 3 & \sqrt{3} & 46 & 3,3 \\ 8 & 26 & 16 & -2 & 0 \\ 10 & 7 & -22 & 4 & 9 \end{pmatrix}, \quad \begin{pmatrix} -2 & \frac{1}{2} \\ 3 & 6 \end{pmatrix}, \quad (9); \quad \begin{pmatrix} -1 \\ 4 \\ -2 \end{pmatrix};$$

$$\begin{pmatrix} 3/4 & 5 & 6 \\ 9 & 67 & 88 \\ -2 & 0 & 9 \end{pmatrix}; \quad \begin{pmatrix} -1 & 12 & \pi & 3 \end{pmatrix}; \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}.$$

In genere le matrici verranno indicate con lettere maiuscole come A, B, eccetera; ad esempio scriveremo:

$$A = \begin{pmatrix} 2 & -1 \\ 3 & -6 \end{pmatrix}.$$

Se una matrice ha m righe e n colonne, si dice che è una *matrice* $m \times n$; ad esempio gli esempi di matrici scritte sopra sono, rispettivamente, 3×5, 2×2, 1×1, 3×1, 3×3, 1×4 e 2×4.

Se gli elementi di una matrice $m \times n$ appartengono a $\mathbb{R}$ (come sarà quasi sempre in questo libro), si dirà che essa appartiene a $\mathbb{R}^{m,n}$ (in molti testi è invece utilizzato, per l'insieme delle matrici $m \times n$, il simbolo $M_{m,n}(\mathbb{R})$). Analogamente si possono definire $\mathbb{Z}^{m,n}$, $\mathbb{Q}^{m,n}$ o $\mathbb{C}^{m,n}$.

Il modo più comune per indicare una generica matrice $A \in \mathbb{R}^{m,n}$ e i suoi elementi è il seguente:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix},$$

ove l'elemento $a_{i,j}$ è quello *di posto* (i, j), cioè che sta sulla i-esima riga e sulla j-esima colonna. Spesso si scrive solo, in modo più compatto:

$$A = (a_{ij})_{i=1,\ldots,m, j=1,\ldots,n},$$

o anche solo $A = (a_{i,j})$, se è noto che la matrice sta in $\mathbb{R}^{m,n}$.

La matrice che ha tutti gli $a_{ij} = 0$ si dice *matrice nulla*; la indicheremo con $O_{m,n} \in \mathbb{R}^{m,n}$, o anche semplicemente O quando sia già inteso a quale insieme $\mathbb{R}^{m,n}$ essa appartenga.

In $\mathbb{R}^{m,n}$ è definita un'operazione di somma nel modo più ovvio (termine a termine):

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1n} \\ b_{21} & b_{22} & 23 & \cdots & 2n \\ \vdots & \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & b_{m3} & \cdots & b_{mn} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & a_{m3} + b_{m3} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

Con questa operazione $(\mathbb{R}^{m,n}, +)$ diviene una struttura algebrica, e si può facilmente verificare che essa è un *gruppo commutativo*.

Proposizione 3.2.1. *La struttura algebrica $(\mathbb{R}^{m,n}, +)$ è un gruppo commutativo.*

Dimostrazione. Dimostriamo che valgono per $(\mathbb{R}^{m,n}, +)$ le proprietà di un gruppo commutativo (vedi Capitolo 1).

1. *La somma è associativa:*

$$\forall A, B, C \in \mathbb{R}^{m,n}, \quad A + (B + C) = A + (B + C).$$

Ciò si ricava immediatamente dal fatto che la somma è associativa in $\mathbb{R}$ e che la somma di matrici si esegue termine a termine.

2. *Esiste l'elemento neutro:* infatti la matrice nulla O, avendo solo elementi nulli, verifica banalmente:

$$O + A = A + O = A, \qquad \forall A \in \mathbb{R}^{m,n}.$$

3. $\forall A \in \mathbb{R}^{m,n}$, *esiste la sua inversa rispetto all'operazione* " + ", *infatti:*

$$(-A) + A = A + (-A) = O,$$

ove se $A = (a_{ij})$, si pone: $-A = (-a_{ij})$.

4. *Commutatività :* Come per il punto 1), è conseguenza immediata della commutatività della somma in $\mathbb{R}$ che si abbia anche

$$\forall A, B \in \mathbb{R}^{m,n}; \quad A + B = B + A.$$

□

Esercizio 3.2.1. Eseguire la verifica dettagliata di 1) e 4) nella dimostrazione sopra.

Un'altra operazione che possiamo fare con le matrici è quella di *moltiplicarle per un numero reale*; questa operazione è una funzione $\mathbb{R} \times \mathbb{R}^{m,n} \to \mathbb{R}^{m,n}$, ove $\forall \alpha \in \mathbb{R}$, $\forall A \in \mathbb{R}^{m,n}$, $\alpha \cdot A$ è definita nel modo seguente:

$$\alpha \cdot A = \alpha \cdot \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{pmatrix}.$$

Avremo (la verifica è piuttosto immediata ed è lasciata per esercizio) che:

Proposizione 3.2.2. [Proprietà del prodotto scalari-matrici] *Qualunque siano $A, B \in \mathbb{R}^{m,n}$ e $\alpha, \beta \in \mathbb{R}$, si ha:*

1. $\alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$;
2. $(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$;
3. $(\alpha\beta) \cdot A = \alpha \cdot (\beta A)$;
4. $1 \cdot A = A$.

Osserviamo che dalle proprietà di cui sopra è facile dedurre che si ha anche:

$$\forall A \in \mathbb{R}^{m,n}, \quad 0 \cdot A = O_{mn}.$$

Infatti $0 \cdot A = (1 - 1) \cdot A = 1 \cdot A + (-1) \cdot A = A + (-A) = O.$

Osservazione 3.2.1. In analogia con quello che si fa con la moltiplicazione numerica nel calcolo letterale, anche nel caso del prodotto tra matrici ometteremo spesso il simbolo dell'operazione, scrivendo semplimente αA piuttosto che $\alpha \cdot A$.

Proposizione 3.2.3. [Insiemi di matrici come spazi vettoriali] *Con le operazioni considerate, si ha che $\forall m, n \in \mathbb{N}^+$, la struttura algebrica $(\mathbb{R}^{m,n}, +, \cdot)$ è uno spazio vettoriale.*

Dimostrazione. Le Proposizioni 3.2.1 e 3.2.2 permettono di affermare che la struttura $(\mathbb{R}^{m,n}, +, \cdot)$ soddisfa la definizione di spazio vettoriale data nel capitolo precedente. □

Esempio 3.2.1. Consideriamo lo spazio vettoriale $\mathbb{R}^{2,3}$ e in esso l'insieme:

$$\mathcal{B} = \{E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23}\};$$

ove E_{ij} è la matrice avente tutti gli elementi nulli tranne $a_{ij} = 1$. Allora avremo che $\mathcal{B}$ è un insieme di generatori per $\mathbb{R}^{2,3}$, infatti comunque si scelga $A \in \mathbb{R}^{2,3}$ si avrà :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = a_{11}E_{11} + a_{12}E_{12} + \cdots + a_{23}E_{23}.$$

Per quanto $(\mathbb{R}^{m,n}, +, \cdot)$ sia uno spazio vettoriale, non utilizzeremo tanto le matrici da questo punto di vista, ma principalmente come strumenti per studiare sistemi lineari di equazioni (vedi Capitolo 4) e funzioni fra spazi vettoriali (vedi i Capitoli 5 e 7).

Notiamo anche che la differenza fra gli spazi vettoriali $\mathbb{R}^{1,n}$ e $\mathbb{R}^n$ è veramente minima; nel primo caso si hanno matrici fatte di una riga con n elementi, nel secondo n-ple ordinate (quindi con le virgole fra un elemento e l'altro). Inoltre le operazioni di somma e prodotto per gli scalari si corrispondono perfettamente. Spesso nel seguito identificheremo implicitamente questi spazi vettoriali (spesso anche gli spazi $\mathbb{R}^n$ e $\mathbb{R}^{n,1}$).

Vediamo adesso alcuni particolari tipi di matrici:

Definizione 3.2.1.

1. *Matrici quadrate.* Quando $n = m$, una matrice $n \times n$ si dice *quadrata*. Esempi:

$$\begin{pmatrix} 2 & 1/2 \\ 3 & -6 \end{pmatrix}, \quad (4), \quad \begin{pmatrix} -2 & 0 & \mathbf{6} \\ 4 & 5 & -6 \\ 7 & 88 & 1 \end{pmatrix}.$$

2. *Matrici triangolari.* Gli elementi di posto (i, i) in una matrice costituiscono la sua *diagonale principale*. Una matrice che abbia solo elementi nulli sotto la sua diagonale principale si dice *triangolare alta* (o *triangolare superiore*); una che li abbia tutti nulli sopra alla diagonale principale si dice *triangolare bassa* (o *triangolare inferiore*).

   Esempi. Triangolari alte (in neretto la diagonale principale):

$$\begin{pmatrix} \mathbf{-2} & 1 & 1 \\ 0 & \mathbf{5} & 0 \\ 0 & 0 & \mathbf{1} \end{pmatrix}; \quad \begin{pmatrix} \mathbf{13} & 21 \\ 0 & \mathbf{-5} \end{pmatrix}; \quad (\mathbf{4}); \quad \begin{pmatrix} \mathbf{6} & 5 & 4 & 3 \\ 0 & \mathbf{1} & -1 & 0 \\ 0 & 0 & \mathbf{3} & 3 \end{pmatrix}.$$

   Triangolari basse:

$$\begin{pmatrix} \mathbf{13} & 0 \\ 7 & \mathbf{-5} \end{pmatrix}; \quad (\mathbf{4}); \quad \begin{pmatrix} \mathbf{1} & 0 & 0 \\ -2 & \mathbf{5} & 0 \\ 3 & 3 & \mathbf{3} \\ 1 & 2 & 3 \end{pmatrix};$$

   Poiché quelle che si usano più spesso sono le triangolari alte le indicheremo semplicemente come *triangolari*.

3. *Matrici diagonali.* Se una matrice è contemporaneamente quadrata, triangolare alta e triangolare bassa, si dice una *matrice diagonale*. In pratica, una matrice diagonale è una matrice quadrata $D = (d_{ij})$, ove ogni elemento fuori dalla diagonale deve essere nullo (cioè $d_{ij} = 0$ se $i \neq j$). Esempi.

$$\begin{pmatrix} \mathbf{13} & 0 \\ 0 & \mathbf{-5} \end{pmatrix}; \quad (\mathbf{5}); \quad \begin{pmatrix} \mathbf{-2} & 0 & 0 \\ 0 & \mathbf{5} & 0 \\ 0 & 0 & \mathbf{1} \end{pmatrix}; \quad \begin{pmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{0} \end{pmatrix}.$$

Useremo talvolta la notazione $D_{a_{11},a_{22},\ldots,a_{nn}}$ per una matrice diagonale $n \times n$, indicando i suoi elementi sulla diagonale principale. Con questa notazione le tre matrici dell'esempio precedente si indicheranno:

$$D_{13,-5}; \quad D_5; \quad D_{-2,5,1}; \quad D_{1,0}.$$

4. *Trasposta di una matrice.* Per ogni $A = (a_{ij}) \in \mathbb{R}^{m,n}$, si definisce *matrice trasposta* di A la matrice ${}^tA = (a_{ji}) \in \mathbb{R}^{n,m}$. In pratica la trasposta di A ha per righe le colonne di A e viceversa. Si può pensare alla trasposta come alla matrice ottenuta scambiando gli elementi di A simmetricamente rispetto alla diagonale principale. Esempi:

$$A = \begin{pmatrix} \mathbf{13} & 0 & 1 \\ 2 & \mathbf{-5} & 2 \end{pmatrix}; \quad {}^tA = \begin{pmatrix} \mathbf{13} & 2 \\ 0 & \mathbf{-5} \\ 1 & 2 \end{pmatrix};$$

$$A = \begin{pmatrix} \mathbf{a} & b \\ c & \mathbf{d} \end{pmatrix}; \quad {}^tA = \begin{pmatrix} \mathbf{a} & c \\ b & \mathbf{d} \end{pmatrix}.$$

5. *Matrici simmetriche e antisimmetriche.* Se una matrice A è tale che $A = {}^tA$, allora A si dice *simmetrica*. Notiamo che una tale matrice è necessariamente quadrata. In una matrice simmetrica si avrà : $a_{ij} = a_{ji}, \forall i, j = \mathbf{1}, \ldots, n$. Esempi:

$$A = \begin{pmatrix} \mathbf{13} & 5 \\ 5 & \mathbf{-5} \end{pmatrix}; \quad (\mathbf{6}); \quad \begin{pmatrix} \mathbf{8} & 2 & 7 \\ 2 & \mathbf{2} & 1 \\ 7 & 1 & \mathbf{9} \end{pmatrix}; \quad \begin{pmatrix} \mathbf{a} & b \\ b & \mathbf{d} \end{pmatrix}.$$

Se una matrice A è tale che $A = -{}^tA$, allora A si dice *antisimmetrica*. Notiamo che una tale matrice è necessariamente quadrata. In una matrice antisimmetrica si avrà : $a_{ij} = -a_{ji}, \forall i, j = \mathbf{1}, \ldots, n$, e quindi $a_{ii} = 0, \forall i = 1, \ldots, n$. Esempi:

$$A = \begin{pmatrix} \mathbf{0} & 5 \\ -5 & \mathbf{0} \end{pmatrix}; \quad (\mathbf{6}); \quad \begin{pmatrix} \mathbf{0} & 2 & 7 \\ -2 & \mathbf{0} & 1 \\ -7 & -1 & \mathbf{0} \end{pmatrix}; \quad \begin{pmatrix} \mathbf{0} & b \\ -b & \mathbf{0} \end{pmatrix}.$$

Vogliamo adesso definire due spazi vettoriali naturalmente associati a ogni matrice, più precisamente, associati alle sue righe e alle sue colonne. Osserviamo che ogni riga di una matrice $m \times n$ può essere considerata come un vettore di $\mathbb{R}^n$ e ogni sua colonna come un vettore di $\mathbb{R}^m$ (ma scritto in verticale invece che nel modo solito).

Ad esempio, se $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, allora le sue righe sono associate ai vettori $\mathbf{R_1} = (1,2,3), \mathbf{R}_2 = (4,5,6) \in \mathbb{R}^3$, e le sue colonne ai vettori $\mathbf{C}_1 = (1,4), \mathbf{C}_2 = (2,5), \mathbf{C}_3 = (3,6) \in \mathbb{R}^2$. Spesso useremo i termini *vettori riga* e *vettori colonna* per indicare tali vettori.

Definizione 3.2.2. Sia $A \in \mathbb{R}^{m,n}$; siano $\mathbf{R}_1, \ldots, \mathbf{R}_m \in \mathbb{R}^n$ i vettori dati dalle sue righe e $\mathbf{C}_1, \ldots, \mathbf{C}_n \in \mathbb{R}^m$ quelli dati dalle sue colonne. Allora:

- il sottostazio vettoriale $\mathcal{R}(A) = \langle \mathbf{R}_1, \ldots, \mathbf{R}_m \rangle \subseteq \mathbb{R}^n$ definito dalle righe di A si dice *Spazio delle righe* di A;
- il sottostazio vettoriale $\mathcal{C}(A) = \langle \mathbf{C}_1, \ldots, \mathbf{C}_n \rangle \subseteq \mathbb{R}^m$ definito dalle colonne di A si dice *Spazio delle colonne* di A.

Quando è chiaro a quale matrice ci stiamo riferendo, scriveremo semplicemente $\mathcal{R}$ o $\mathcal{C}$.

Esempio 3.2.2. Sia $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 0 \end{pmatrix} \in \mathbb{R}^{2,3}$.

In questo caso lo spazio delle righe è

$$\mathcal{R}(A) = \langle (1,2,3), (1,-1,0) \rangle \subseteq \mathbb{R}^3,$$

mentre quello delle colonne è

$$\mathcal{C}(A) = \langle (1,1), (2,-1), (3,0) \rangle \subseteq \mathbb{R}^2.$$

Notiamo che $\dim \mathcal{R}(A) = 2$ (poiché i vettori $(1,2,3), (1,-1,0)$ sono linearmente indipendenti), e $\dim \mathcal{C}(A) = 2$, cioè $\mathcal{C} = \mathbb{R}^2$, infatti se prendiamo due qualunque dei tre vettori colonna, essi saranno linearmente indipendenti (mentre non possono essere tutti e tre linearmente indipendenti perché siamo in $\mathbb{R}^2$, quindi $\dim \mathcal{C}(A) \leq 2$).

Esercizio 3.2.2. Data $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$, determinare $\dim \mathcal{R}(A)$ e $\dim \mathcal{C}(A)$.

3.3 Il prodotto fra matrici

Meno ovvia è invece l'operazione di *prodotto fra matrici*, che però sarà quella su cui più si appunterà il nostro interesse. Innanzi tutto questa non è in generale un'operazione fra matrici della stessa forma, ma è un'applicazione:

$$\mathbb{R}^{m,n} \times \mathbb{R}^{n,t} \to \mathbb{R}^{m,t}. \tag{3.1}$$

Cioè potremo moltiplicare A per B solo se il numero di colonne di A coincide col numero delle righe di B, e il risultato sarà una matrice con il numero di righe di A e il numero di colonne di B.

Vediamo come si definisce il prodotto nel caso più semplice, quello di una matrice $1 \times n$ con una $n \times 1$.

$$(a_{11}a_{12}\dots a_{1n}) \cdot \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix} = (a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} + \dots + a_{1n}b_{n1}) \in \mathbb{R}^{1,1}$$

il risultato è una matrice 1×1, e quindi corrisponde a un elemento di $\mathbb{R}$, a un numero reale.

Esempio 3.3.1.

$$\begin{pmatrix} 2 & 4 & 18 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 0 \\ 10 \end{pmatrix} = (2.2 + 4.(-1) + 18.0 + (-1).10) =$$

$$= (4 - 4 - 10) = (-10).$$

In generale, il prodotto di una matrice $A \in \mathbb{R}^{m,n}$ per una matrice $B \in \mathbb{R}^{n,t}$, si effettuerà *righe per colonne*; in dettaglio, avremo che se dobbiamo eseguire il prodotto $A \cdot B = C$:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1t} \\ b_{21} & b_{22} & \cdots & b_{2t} \\ b_{31} & b_{32} & \cdots & b_{3t} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nt} \end{pmatrix} =$$

$$= \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1t} \\ c_{21} & c_{22} & \cdots & c_{2t} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mt} \end{pmatrix} = (c_{ij})_{i=1,\dots,m, j=1,\dots,t} \in \mathbb{R}^{m,t},$$

cioè se $C \in \mathbb{R}^{m,t}$ è la matrice risultato del nostro prodotto, si avrà che

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} + \dots + a_{1n}b_{n1}$$

sarà il prodotto della prima riga di A per la prima colonna di B (fatto come mostrato nell'esempio precedente); c_{12} sarà il prodotto della prima riga di A per la seconda colonna di B, e così via.

In generale l'elemento di posto (i, j) in C è dato dal prodotto dell'i-esima riga di A per la j-esima colonna di B:

$$c_{ij} = (a_{i1}\ a_{i2}\ \dots\ a_{in}) \cdot \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} =$$

$$= (a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj}) = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Riassumendo, vediamo la regola per il prodotto di matrici,

Definizione 3.3.1. *Date due matrici $A = (a_{ik}) \in \mathbb{R}^{m,n}$ e $B = (b_{kj}) \in \mathbb{R}^{n,t}$, il loro prodotto è la matrice $C \in \mathbb{R}^{m,t}$ data da:*

$$A \cdot B = C = (c_{ij}), \qquad \text{ove} \quad c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Osserviamo che la definizione di questa operazione, data attraverso doppi indici e sommatorie, può apparire più complicata di quanto effettivamente non sia. Per prendere dimestichezza con l'esecuzione di queste operazioni è bene osservare gli esempi e provare a eseguirne qualcuna da soli (anche inventandosela).

Esempio 3.3.2.

$$\begin{pmatrix} 2 & 1 & 10 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ -1 & 6 \end{pmatrix} =$$

$$= \begin{pmatrix} 2.2 + 1.1 + 10(-1) & 2.0 + 1.1 + 10 \cdot 6 \\ 1.2 + 0.1 + (-1)(-1) & 1.0 + 0.1 + 1(-6) \\ 2.2 + 1.1 + 0 & 2.0 + 1.1 + 0.6 \end{pmatrix} = \begin{pmatrix} -5 & 61 \\ 3 & -6 \\ 5 & 1 \end{pmatrix};$$

$$\begin{pmatrix} 1 & 3 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} =$$

$$\begin{pmatrix} 1.2 + 3.1 & 1.0 + 3.1 & 1.1 + 3 \cdot 2 \\ 0.2 + (-1).1 & 0.0 + (-1).1 & 0.1 + (-1).2 \end{pmatrix} = \begin{pmatrix} 5 & 3 & 7 \\ -1 & -1 & -2 \end{pmatrix}.$$

Proposizione 3.3.1. [Proprietà del prodotto di matrici.] *Date comunque delle matrici* $A, A_1, A_2 \in \mathbb{R}^{m,n}$, $B, B_1, B_2 \in \mathbb{R}^{n,t}$, $C \in \mathbb{R}^{t,u}$, *si ha:*

a) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$. *(Proprietà associativa);*

b) $(A_1 + A_2) \cdot B = A_1 \cdot B + A_2 \cdot B$; *(Proprietà distributive);*
$A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2$

c) $\forall \alpha \in \mathbb{R}, \quad \alpha(A \cdot B) = (\alpha A) \cdot B = A \cdot (\alpha B)$;

d) $O_{mn} \cdot B = O_{mt}$; $\quad A \cdot O_{nt} = O_{mt}$.

Dimostrazione. La dimostrazione segue abbastanza immediatamente da come sono definite le operazioni, ed è lasciata per esercizio al lettore. □

Osservazione 3.3.1. Anche per l'operazione di prodotto fra matrici è comune l'uso di ometterne il simbolo, scrivendo semplicemente AB invece di $A \cdot B$. Nel seguito questa sarà la convenzione che adotteremo più spesso.

Mostriamo adesso dei semplici esempi di applicazioni pratiche del linguaggio delle matrici e di come l'operazione di prodotto che abbiamo definito rispecchi il trattamento che spesso si fa di matrici di dati.

Esempio 3.3.3. Una ditta commercia in Arance, Banane e Ciliegie, che tiene in due magazzini. Nel primo ci sono, rispettivamente, 1, 2 e 1,5 tonnellate di ciascuna merce, mentre nel secondo $3t$ di banane e $2t$ di ciliegie. Il prezzo dei tre tipi di frutta è di 1, 2 e 4 migliaia di euro alla tonnellata. Che valore si ha in ogni magazzino? La situazione delle merci è sintetizzata nel seguente diagramma:

	A	B	C
$Mag.$ 1	1	2	1,5
$Mag.$ 2	0	3	2

Naturalmente (è un esercizio da scuola primaria) la risposta si ha calcolando:

$Mag.\ 1: \ 1 \times 1000 + 2 \times 2000 + 1,5 \times 4000 = 11000;$

$Mag.\ 2: \ 0 \times 1000 + 3 \times 2000 + 2 \times 4000 = 14000.$

Rivediamolo in notazione matriciale. Per ricavare il valore delle merci, si deve moltiplicare la matrice 2×3 (magazzini / frutta) che raccoglie

i dati in tabella, per la matrice 3×1 (frutta / migliaia di euro): $\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$, data dalla colonna dei prezzi, ottenendo:

$$\begin{pmatrix} 1 & 2 & 1,5 \\ 0 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1+4+6 \\ 0+6+8 \end{pmatrix} = \begin{pmatrix} 11 \\ 14 \end{pmatrix}.$$

Dalla matrice 2×1 (magazzini / migliaia di euro) che è il risultato del prodotto si ricava che ci sono 11.000 euro di merce nel magazzino 1, e 14.000 nel magazzino 2.

Esempio 3.3.4. Supponiamo che ogni anno l'80% degli abitanti della Toscana resti nella loro regione, mentre il restante 20% emigri altrove in Italia. Nel resto d'Italia, il 10% degli abitanti emigri in Toscana e il 90% ne resti fuori. Se in un dato anno la popolazione della Toscana è di 7 milioni e quella del resto d'Italia è di 50 milioni, quale sarà la situazione dopo un anno? E dopo due o tre?

Avremo, indicando con T i la popolazione in Toscana, con $I - T$ quella nel resto d'Italia, con $\to T$ le "entrate in Toscana" e con $\to I - T$ le "entrate nel resto d'Italia":

	T	$I-T$
$\to \quad T$	0,8	0,1
$\to I-T$	0,2	0,9

Si possono inserire i dati della tabella in una matrice e per ricavare l'ammontare degli spostamenti di popolazione si deve moltiplicarla per $\begin{pmatrix} 7 \\ 50 \end{pmatrix}$, la matrice che contiene i dati iniziali.

$$\begin{pmatrix} 0,8 & 0,1 \\ 0,2 & 0,9 \end{pmatrix} \begin{pmatrix} 7 \\ 50 \end{pmatrix} = \begin{pmatrix} (0,8).7 + (0,1).50 \\ (0,2).7 + (0,9).50 \end{pmatrix} = \begin{pmatrix} 10,6 \\ 46,4 \end{pmatrix} \begin{matrix} T \\ I-T \end{matrix}.$$

Quindi dopo un anno ci sono 10,6 milioni di persone in Toscana e 46,4 milioni fuori di essa.

Per conoscere i dati dopo due anni si ripete l'operazione usando il vettore $\begin{pmatrix} 10,6 \\ 46,4 \end{pmatrix}$ come dato iniziale, ottenendo:

$$\begin{pmatrix} 0,8 & 0,1 \\ 0,2 & 0,9 \end{pmatrix} \begin{pmatrix} 10,6 \\ 46,4 \end{pmatrix} = \begin{pmatrix} 13,2 \\ 35,88 \end{pmatrix} \begin{matrix} T \\ I-T \end{matrix}.$$

Esempio 3.3.5. Nella situazione dell'esempio precedente, esiste una distribuzione di popolazione iniziale tale che alla fine dell'anno quei flussi migratori non cambieranno la situazione?

Sia x la popolazione toscana e y quella del resto d'Italia (in milioni) all'inizio dell'anno. Si dovrà avere che con i flussi dati le due popolazioni alla fine dell'anno siano invariate, quindi:

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

cioè

$$\begin{pmatrix} 0,8 & 0,1 \\ 0,2 & 0,9 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0,8x + 0,1y \\ 0,2x + 0,9y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

Questa *equazione matriciale*, più la condizione che $x + y = 57$ (se vogliamo usare lo stesso dato per la popolazione italiana complessiva) equivalgono al sistema

$$\begin{cases} 0,8x + 0,1y = x \\ 0,2x + 0,9y = y \\ x + y = 57 \end{cases}.$$

Che ha come soluzione:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 19 \\ 38 \end{pmatrix}.$$

Quindi se la situazione fosse che ci sono 19 milioni di abitanti in Toscana e 38 nel resto d'Italia, con i flussi migratori descritti tale distribuzione rimarrebbe invariata.

Esempio 3.3.6. [Studio di una catena alimentare] Supponiamo di studiare un ecosistema chiuso, ad esempio dato dalle principali specie viventi un un piccolo ambiente isolato. Le specie da considerare siano:

1) r specie di piante : $p_1, p_2, \ldots, p_r$;

2) s specie di erbivori : $e_1, e_2, \ldots, e_s$;

3) t specie di carnivori : $c_1, c_2, \ldots, c_t$.

Supponiamo che non ci siano onnivori, cioè gli e_j si nutrano solo delle piante p_i mentre i carnivori c_k solo degli erbivori e_j. Consideriamo poi, in un periodo di un anno, le seguenti quantità :

x_{ij} = l'ammontare (in Kg) della massa di piante p_i mangiate da un e_j;

y_{jk} = il numero di capi della specie e_j mangiati da un membro di c_k.

Allora la matrice $X = (x_{ij})_{i=1,\ldots,r;j=1,\ldots,s}$ descrive il primo ciclo della catena alimentare costituita dal nostro sistema, mentre il secondo ciclo è descritto dalla matrice $Y = (y_{jk})_{j=1,\ldots,s;k=1,\ldots,t}$. Cioè X contiene tutti i dati relativi a come le specie di erbivori dipendono dalle varie specie di piante, e la Y quelli di come i carnivori dipendono dalle varie specie di erbivori.

Se vogliamo sapere come i carnivori dipendono dalle varie specie di piante, dobbiamo risolvere il seguente problema: qual è la quantità z_{ik} di piante p_i necessaria al sostentamento di un c_k (indirettamente, attraverso gli erbivori)?

Partiamo da un c_1: esso avrà mangiato (in un anno) y_{11} capi di e_1, y_{21} capi di e_2, ..., y_{s1} capi di e_s (i dati nella prima colonna di Y). Ma ogni e_1 ha mangiato a sua volta $x_{11}Kg$ di piante p_1, ogni e_2 ne ha mangiati x_{12} ecc.

Così c_1 ha consumato $z_{11} = (x_{11}y_{11} + x_{12}y_{21} + \cdots + x_{1s}y_{s1})\ Kg$ di p_1, e in generale la quantità z_{ik} di p_i consumata indirettamente da un c_k sarà:

$$z_{ik} = x_{i1}y_{1k} + x_{i2}y_{2k} + \cdots + x_{is}y_{sk}$$

che è proprio l'elemento di posto (i,k) nella matrice prodotto $Z = X \cdot Y \in \mathbb{R}^{r,t}$:

$$Z = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1s} \\ \vdots & \vdots & & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{is} \\ \vdots & \vdots & & \vdots \\ x_{r1} & x_{r2} & \cdots & x_{rs} \end{pmatrix} \begin{pmatrix} y_{11} & \cdots & y_{1k} & \cdots & y_{1t} \\ y_{21} & \cdots & y_{2k} & \cdots & y_{2t} \\ \vdots & & \vdots & & \vdots \\ y_{s1} & \cdots & y_{sk} & \cdots & y_{st} \end{pmatrix}$$

Quindi la matrice $Z = (\ z_{ik}\ )_{i=1,\ldots,r;k=1,\ldots,t}$ contiene tutti i dati sull'assunzione delle specie vegetali da parte dei carnivori.

3.4 L'anello delle matrici quadrate

In generale abbiamo visto che il prodotto si fa fra matrici di forma diversa, quindi non è un'operazione in $\mathbb{R}^{m,n}$, tranne nel caso in cui nella (3.1) si abbia $m = n = t$. In questo caso otterremo che il prodotto tra matrici è una funzione $\mathbb{R}^{n,n} \times \mathbb{R}^{n,n} \to \mathbb{R}^{n,n}$, cioè è un'operazione in $\mathbb{R}^{n,n}$ e quindi si ha una struttura algebrica $(\mathbb{R}^{n,n}, \cdot)$. Vediamo quali siano le sue proprietà principali.

- *In* $(\mathbb{R}^{n,n}, \cdot)$ *esiste l'elemento neutro.* Vediamolo prima nel caso 2×2; $\forall\, a, b, c, d$ si ha:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

  Quindi la matrice diagonale $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ si comporta come elemento neutro rispetto al prodotto in $\mathbb{R}^{2,2}$.

  Il caso generale $n \times n$ è del tutto analogo, e a fungere da elemento neutro sarà la matrice diagonale:

$$I_n = D_{1,\ldots,1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{ij})_{i,j=1,\ldots,n}$$

ove $\delta_{ij} = \begin{cases} 0 \text{ se } i \neq j \\ 1 \text{ se } i = j \end{cases}$ (δ_{ij} è detto *simbolo di Kroneker*).

Avremo infatti che se $AI_n = C = (c_{ij})$, si ha

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\delta_{kj} = a_{i1}\delta_{1j} + a_{i2}\delta_{2j} + \cdots + a_{ij}\delta_{jj} + \cdots + a_{in}\delta_{nj},$$

ma $\delta_{kj} = 0$ per $k \neq j$, mentre $\delta_{jj} = 1$, cosicché nella sommatoria rimarrà soltanto $c_{ij} = a_{ij}\delta_{jj} = a_{ij}$. Questo avviene per ogni i, j, e quindi si otterrà $C = A$.

Un computo analogo mostra che $I_n A = A$.

In conclusione, $\forall n \in \mathbb{N}^+$, I_n è l'elemento neutro rispetto al prodotto di matrici in $(\mathbb{R}^{n,n}, \cdot)$.

La matrice I_n si dice *matrice identità* e si scrive semplicemente I quando non ci sia necessità di specificare n.

- *Non vale la proprietà commutativa.* Ad esempio:

$$\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 8 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 \\ -1 & 6 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 11 & 6 \end{pmatrix}.$$

- *Non sempre esiste l'inverso di una matrice rispetto al prodotto.*

  Consideriamo la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2,2}$. L'elemento inverso di A in $(\mathbb{R}^{2,2}, \cdot)$ dovrebbe essere una matrice B per la quale si abbia $AB = BA = I$. Si può però vedere che il prodotto

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix},$$

  e, qualunque siano $c, d \in \mathbb{R}$, questa matrice è $\neq I$, perciò non esiste l'inversa della matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

- *Esistono "divisori dello zero".* Esistono matrici non nulle il cui prodotto dia O (in un anello tali elementi vengono detti *divisori di zero*); ad esempio:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 5 & 7 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

  Notiamo che se due matrici A, B sono divisori di zero, non possono ammettere matrice inversa. Se infatti esistesse una matrice A', inversa di A, si avrebbe $O = (BA)A' = B(A \cdot A') = BI = B$, che è falso.

Da quanto appena visto si ha che la struttura $(\mathbb{R}^{n,n}, \cdot)$ **non** è un gruppo. Si ha invece che la struttura $(\mathbb{R}^{n,n}, +, \cdot)$ è un *anello* (vedi Capitolo 1), non commutativo.

Notiamo che ci sono anche matrici (quadrate) che ammettono inversa, ad esempio:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Proposizione 3.4.1. [Unicità dell'inversa] *L'inversa di una matrice A, se esiste, è unica e si indica con A^{-1}.*

Dimostrazione. Infatti se esistessero due matrici A' e A'', entrambe inverse di A, allora si avrebbe:

$$A' = A'I = A'(AA'') = (A'A)A'' = IA'' = A''.$$

e quindi le due matrici coinciderebbero. □

Definizione 3.4.1. *Una matrice $A \in \mathbb{R}^{n,n}$ si dice invertibile, o non singolare, se esiste la sua inversa A^{-1}. Le matrici che non ammettono inversa vengono invece definite singolari.*

Proposizione 3.4.2. [Prodotto di matrici invertibili] *Date due matrici invertibili $A, B \in \mathbb{R}^{n,n}$, anche la matrice prodotto AB risulta invertibile, e la sua inversa è: $(AB)^{-1} = B^{-1}A^{-1}$.*

Dimostrazione. Basta svolgere (usando la proprietà associativa) il prodotto: $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$. □

Osservazione 3.4.1. *Potenze di matrici quadrate.* Per le matrici quadrate, possiamo considerare l'operazione di *potenza* di una matrice; ad esempio per il quadrato di una matrice si definisce: $A^2 = AA$. Si potrebbe pensare che sia più problematico definire A^3, infatti se definiamo $A^3 = A^2A$, chi ci garantisce che questo sia uguale a AA^2, visto

che il prodotto di matrici non è commutativo? Ma in effetti il problema non sussiste; grazie alla proprietà associativa, avremo infatti: $A^2A = (AA)A = A(AA) = AA^2$. Si può quindi generalizzare il concetto di potenza di una matrice (quadrata) in quanto la proprietà associativa del prodotto ci garantisce che non è rilevante in che ordine si eseguano i vari prodotti dei fattori A.

Definizione 3.4.2. *Data $A \in \mathbb{R}^{n,n}$, si definisce, per ogni intero $k \geq 2$, la matrice potenza $A^k \in \mathbb{R}^{n,n}$ come il prodotto di k fattori consecutivi uguali ad A:*

$$A^k = AA \cdots A \qquad (k \text{ volte}).$$

In analogia con quanto si fa con le potenze numeriche, porremo anche:

$$A^0 = I \quad e \quad A^1 = A.$$

Esempio 3.4.1. Riprendiamo la situazione descritta nell'Esempio 3.3.4; avevamo la matrice $A = \begin{pmatrix} 0,8 & 0,1 \\ 0,2 & 0,9 \end{pmatrix}$, che rappresentava dei flussi migratori annuali da e per la Toscana. Ammettendo che l'andamento di tali flussi rimanga lo stesso negli anni, per conoscere i dati dopo due anni si ripeteva l'operazione di prodotto sulla matrice dei dati iniziali, calcolando:

$$A\left(A\begin{pmatrix} 7 \\ 50 \end{pmatrix}\right) = A\begin{pmatrix} 10,6 \\ 46,4 \end{pmatrix}$$

Ma si può anche considerare:

$$A\left(A\begin{pmatrix} 7 \\ 50 \end{pmatrix}\right) = (AA)\begin{pmatrix} 7 \\ 50 \end{pmatrix} = A^2\begin{pmatrix} 7 \\ 50 \end{pmatrix}.$$

Ciò mostra che la matrice A^2 descrive i flussi migratori *biennali*, e in generale, calcolando la matrice A^k, si ha che la moltiplicazione $A^k\begin{pmatrix} 7 \\ 50 \end{pmatrix}$ ci darà la situazione dopo k anni, cioè la matrice A^k rappresenta il flusso migratorio in k anni.

3.5 Esercizi

Esercizio 3.5.1. Eseguire i seguenti prodotti:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = ? \qquad \begin{pmatrix} 7 \\ 4 \end{pmatrix}\begin{pmatrix} 2 & -1 \end{pmatrix} = ?$$

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix}\begin{pmatrix} \frac{a}{a^2-b^2} & \frac{-b}{a^2-b^2} \\ \frac{-b}{a^2-b^2} & \frac{a}{a^2-b^2} \end{pmatrix} = ?$$

Esercizio 3.5.2. Eseguire i seguenti prodotti:

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \\ 0 & 1 \end{pmatrix} = ? \qquad \begin{pmatrix} 1 & -1 \\ -1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 4 \end{pmatrix} = ?$$

Esercizio 3.5.3. Siano date le seguenti matrici:

$$A = (1 \quad 0 \quad -1), \; B = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}.$$

Eseguire i prodotti AB e BA se possibile.

Esercizio 3.5.4. Quali condizioni devono soddisfare le matrici A, B, C affinché il seguente prodotto sia eseguibile: $A(BC)$?

Esercizio 3.5.5. Dal terzo prodotto dell'Esercizio 3.5.1, si può dedurre che ogni matrice simmetrica in $\mathbb{R}^{2,2}$ è invertibile?

Esercizio 3.5.6. Risolvere la seguente equazione matriciale: $AX = B$, cioè determinare la matrice incognita X che verifica l'uguaglianza, ove:

$$A = \begin{pmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 7 \end{pmatrix} \in \mathbb{R}^{3,2}, \; X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2,1}; \; B = \begin{pmatrix} 2 \\ 1 \\ 6 \end{pmatrix} \in \mathbb{R}^{3,1}.$$

Esercizio 3.5.7. Dimostrare che se una matrice quadrata $A \in \mathbb{R}^{n,n}$ è tale che $A^2 = I$, allora A è uguale alla propria inversa.

Esercizio 3.5.8. Verificare che la matrice $A = \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix} \in \mathbb{R}^{2,2}$ è soluzione dell'equazione matriciale $A^2 - A - 8I = O$.

Esercizio 3.5.9. Siano $A \in \mathbb{R}^{m,n}$, $B \in \mathbb{R}^{n,m}$. Verificare che da $AB = O$ e $BA = O$ non si può dedurre che $AB = BA$.

Esercizio 3.5.10. Stabilire se esistono valori del parametro reale k tale che la matrice $A = \begin{pmatrix} 1 & 0 & k \\ k & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ sia idempotente (cioè che $A^2 = A$).

Esercizio 3.5.11. Verificare che se $AB = A$ e $BA = B$ allora A è idempotente.

Esercizio 3.5.12. Dimostrare che se $A \in \mathbb{R}^{n,n}$ allora $A + {}^tA$ è una matrice simmetrica.

Esercizio 3.5.13. Considerare l'insieme $\mathcal{T}^a_{n,n}$ delle matrici triangolari alte in $\mathbb{R}^{n,n}$ e quello $\mathcal{D}_{n,n}$ delle matrici diagonali in $\mathbb{R}^{n,n}$. Dimostrare che $(\mathcal{T}^a_{n,n}, +, \cdot)$ e $(\mathcal{D}_{n,n}, +, \cdot)$ sono anelli, e il secondo è commutativo.

Esercizio 3.5.14. Supponiamo di voler acquistare n divesi tipi di merci, diciamo b_j, $j = 1, \ldots, n$ unità di ciascuna di esse. Abbiamo a disposizione m negozi ove effettuare l'acquisto, e sia a_{ij} il prezzo unitario della j-esima merce effettuato nell'i-esimo negozio. Con quale operazione matriciale possiamo ottenere i costi totali $c_1, \ldots, c_m$ della nostra spesa a seconda del negozio scelto?

Osserviamo che il minimo tra i c_i ci dirà in quale negozio sia più conveniente effettuare l'acquisto.

Esercizio 3.5.15. * Un giardiniere di Keukenhof deve programmare le aiuole di tulipani per l'anno venturo. Ha deciso di combinare tre specie diverse a, b, c di tulipani nella medesima aiuola. Della specie a possiede in magazzino x_a bulbi, della specie b ne possiede x_b e della specie c, x_c. Di tutte e tre le specie ne aveva seppelliti y_a, y_b, y_c l'anno precedente, esse a fine stagione daranno approssimativamente $2y_a, 6y_b, 8y_c$ bulbi ciascuna. Supponiamo che il giardiniere voglia utilizzare $2/3$ dei bulbi in magazzino e $1/4$ dei bulbi che avrà a fine stagione. Supponiamo inoltre che le proporzioni con cui voglia utilizzare i bulbi a, b, c siano l'una il doppio dell'altra (ossia i bulbi c il doppio dei bulbi b, e i bulbi b il doppio dei bulbi a). Quale operazione matriciale deve impostare il giardiniere per sapere quanti bulbi dovrà usare l'anno prossimo per ciascuna specie?

Esercizio 3.5.16. Considerare l'insieme di generatori $\mathcal{B} \subset \mathbb{R}^{2,3}$ che appare nell'Esempio 3.2.1 e dimostrare che è una base.

Esercizio 3.5.17. In generale, qual è la dimensione dello spazio vettoriale $\mathbb{R}^{m,n}$?

Esercizio 3.5.18. Stabilire se una delle seguenti matrici può essere scritta come combinazione lineare delle altre due:

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, C = \begin{pmatrix} 4 & 1 \\ 8 & 1 \end{pmatrix}.$$

Esercizio 3.5.19. Considerare lo spazio vettoriale $(\mathbb{R}^{2,3}, +, \cdot)$ e trovarne una base. Più in generale, qual è la dimensione dello spazio vettoriale $\mathbb{R}^{m,n}$?

Esercizio 3.5.20. Date le matrici quadrate A e B dello stesso ordine, è vero che $A^2 - B^2 = (A - B)(A + B)$?

Esercizio 3.5.21. Siano

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},\ B = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix},\ C = \begin{pmatrix} 0 & 2 \\ -2 & 1 \end{pmatrix}.$$

Calcolare: $2A - B$, $3A + 2B - 4C$, $-2A + B + 2C - 2B$, $3B + 2(2A - C) - (A + B + 2C)$.

Risolvere, se possibile:
$3X + 2(A - X) + B + 2(C + 2X) = 0$,
$4A + 2(B + 2X) - 3(C + X + 2A) = 0$,
$4(A + B + X) + 4(-A - B + X) - 4(A - B + X) = 0$.

Esercizio 3.5.22. Siano

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},\ B = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix},\ C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \\ 1 & -2 \end{pmatrix}.$$

Calcolare, quando possibile: AB, AC, BA, BC, $({}^tA)B$, $({}^tC)A$.

Esercizio 3.5.23. Calcolare, se possibile, i seguenti prodotti: AB e BA dove: $A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ -1 & 0 \\ 5 & 7 \end{pmatrix}$.

Esercizio 3.5.24. Verificare che $A(BC) = (AB)C$, ove A, B, C siano:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix},\quad B = \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 2 \end{pmatrix},\quad C = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}.$$

Esercizio 3.5.25. Mostrare che due matrici quadrate diagonali dello stesso ordine commutano sempre, ossia per ogni $A = (a_{i,j}), B = (b_{i,j}) \in \mathbb{R}^{n,n}$ con $a_{i,j} = b_{i,j} = 0$ se $i \neq j$ si ha che $AB = BA$.

Esercizio 3.5.26. Sia $A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$ con $a, b, c \in \mathbb{R}$ tutti distinti. Determinare tutte le matrici $B \in \mathbb{R}^{3,3}$ che commutano con A, ossia, tali che $AB = BA$.

Esercizio 3.5.27. Data $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2,2}$, consideriamo $B \in \mathbb{R}^{2,2}$, una matrice tale che $AB = BA$.

Dimostrare che esistono $h, k \in \mathbb{R}$ tale che $B = hI + \begin{pmatrix} 0 & k \\ 0 & 0 \end{pmatrix}$.

Esercizio 3.5.28. Sia $A = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ con $a \in \mathbb{R}$. Calcolare A^n per ogni numero naturale n.

Esercizio 3.5.29. Dimostrare che se $A \in \mathbb{R}^{3,3}$ è una matrice che commuta con ogni matrice $B \in \mathbb{R}^{3,3}$, allora necessariamente A è un multiplo dell'identità.

Esercizio 3.5.30. Trovare due matrici quadrate A e B non nulle e diverse tra loro tali che $AB = O$.

Esercizio 3.5.31. Sia data la matrice quadrata $A = \begin{pmatrix} 0 & 1 \\ k & 1 \end{pmatrix}$. Dire, al variare di k, se esiste una matrice B tale che $AB = I$.

Esercizio 3.5.32. È vero che se A è una matrice di rango massimo allora anche $A + {}^tA$ ha rango massimo?

Esercizio 3.5.33. * Ricordiamo che una matrice si dice *idempotente* se $A^2 = A$. Dimostrare che se A e B sono due matrici tali che $AB = A$ e $BA = B$ allora A e B sono idempotenti.

Esercizio 3.5.34. Sia A una matrice $m \times n$ e sia B una matrice $n \times p$. Dimostrare che ${}^t(AB) = ({}^tB)({}^tA)$.

Esercizio 3.5.35. Mostrare che per ogni coppia di matrici vale la seguente uguaglianza:

$${}^t(A + B) = {}^t A + {}^t B.$$

Esercizio 3.5.36. * Sia data la matrice quadrata $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Determinare, al variare di n, la matrice A^n. Determinare, se possibile, la matrice A^{-1} e, se tale matrice esiste, verificare che $(A^n)^{-1} = (A^{-1})^n$.

Esercizio 3.5.37. Date due matrici quadrate A e B dello stesso ordine, è vero che $A^2 - B^2 = (A - B)(A + B)$?

Esercizio 3.5.38. Fare un esempio di due matrici quadrate $A, B \in \mathbb{R}^{3,3}$ tali che $AB \neq BA$.

Esercizio 3.5.39. * Sia data la matrice quadrata $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Determinare, al variare di n, la matrice A^n. Determinare, se possibile, la matrice A^{-1} e, se tale matrice esiste, verificare che $(A^n)^{-1} = (A^{-1})^n$.

Esercizio 3.5.40. Una matrice quadrata $A = (a_{i,j})_{i,j=1,\dots,n}$ si dice antisimmetrica se $a_{i,j} = -a_{j,i}$ per ogni $i = 1, \dots, n$. Dimostrare che:

1. Gli elementi sulla diagonale principale di una matrice antisimmetrica sono tutti nulli;
2. A è antisimmetrica se e solo se ${}^tA = -A$.

CAPITOLO 4

Sistemi lineari

Prerequisiti: Matrici.

Obiettivi: Saper risolvere sistemi lineari di equazioni; conoscere l'algoritmo di eliminazione di Gauss per trovare il rango di una matrice; conoscere il teorema di Rouché–Capelli per sapere se un sistema ha soluzioni o meno (e quante ne ha) senza risolverlo.

4.1 Introduzione

Partiamo considerando il seguente modello economico, detto modello *input-output*, dovuto all'economista russo naturalizzato statunitense Wassily Leontief (1906 - 1999; premio Nobel 1973 per l'economia). Consideriamo un'economia fatta di n industrie $I_1, \dots, I_n$ e una popolazione di consumatori. Per ogni $i \in \{1, \dots, n\}$, sia x_i il valore di merce prodotta dall'industria I_i, e b_i quello della domanda di quella merce da parte dei consumatori (x_i e b_i saranno espressi nella valuta corrente, ad esempio in euro o dollari). Non tutta la produzione industriale sarà disponibile ai consumatori, perché una parte sarà consumata dalle industrie stesse nel loro processo di produzione: sia x_{ij} il valore dell'output di I_i consumato dall'industria I_j. Assumiamo che x_{ij} sia proporzionale a x_j, diciamo $x_{ij} = a_{ij}x_j$ (è un'ipotesi semplificatoria, ma abbastanza plausibile). Consideriamo la quantità

$$x_i - \sum_{j=1}^{n} x_{ij} = x_i - \sum_{j=1}^{n} a_{ij}x_j$$

questa rappresenta il valore della produzione dell'industria I_i che sarà disponibile per i consumatori.

L'economia in discussione sarà in equilibrio se la quantità di ogni merce prodotta disponibile sul mercato sarà pari alla richiesta di tale merce, cioè se si verificano le uguaglianze:

$b_1 = x_1 - \sum_{j=1}^{n} a_{1j}x_j$;

...

$b_n = x_n - \sum_{j=1}^{n} a_{nj}x_j$.

Assumendo che i coefficienti a_{ij} e le quantità b_i siano noti, la questione sarà :

Quanto deve produrre ogni industria I_i affinchè l'economia sia in equilibrio?

Come si vede, la soluzione a tale problema (se esiste, un punto di equilibrio potrebbe anche non esistere) sarà data dalle soluzioni del sistema:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

Questo sarà il problema che considereremo in questa sezione: risolvere i sistemi lineari di equazioni (nel campo reale $\mathbb{R}$).

Esempio 4.1.1. Consideriamo il seguente sistema:

$$\begin{cases} 3x - 2y = 1 \\ x + y + z = 3 \\ 2x - 3y - z + 2 = 0 \end{cases}.$$

Risolvere questo sistema di equazioni nelle incognite x, y, z significa trovare tutte le terne (x, y, z) di numeri reali che soddisfano le equazioni date. Possiamo ad esempio procedere "per sostituzione" partendo dalla prima equazione (che contiene solo due variabili), ricavandoci la x e sostituendo nelle altre:

$$\begin{cases} x = \frac{1}{3}(2y+1) \\ \frac{1}{3}(2y+1) + y + z = 3 \\ \frac{2}{3}(2y+1) - 3y - z + 2 = 0 \end{cases} \rightarrow \begin{cases} x = \frac{1}{3}(2y+1) \\ (2y+1) + 3y + 3z = 9 \\ (4y+2) - 9y - 3z + 6 = 0 \end{cases}.$$

Da cui

$$\begin{cases} x = \frac{1}{3}(2y+1) \\ 5y + 3z = 8 \\ -5y - 3z = -8 \end{cases} \rightarrow \begin{cases} x = \frac{1}{3}(2y+1) \\ y = \frac{-3z+8}{5} \\ -5(\frac{-3z+8}{5}) - 3z = -8 \end{cases},$$

$$\begin{cases} x = \frac{2}{3}(\frac{-3z+8}{5}) + \frac{1}{3} \\ y = \frac{-3z+8}{5} \\ -8 = -8 \end{cases} \rightarrow \begin{cases} x = \frac{-2z+7}{5} \\ y = \frac{-3z+8}{5} \\ -8 = -8 \end{cases}.$$

Osserviamo che la terza equazione è divenuta un'identità e quindi non impone nessuna condizione alle variabili, che la z è rimasta libera di

assumere qualsiasi valore, mentre la x e la y rimangono assegnate una volta fissata z. Quindi il sistema ha infinite soluzioni date da tutte le terne del tipo:

$$(x, y, z) = \left(\frac{7-2z}{5}, \frac{8-3z}{5}, z\right), \quad \forall z \in \mathbb{R}.$$

In questo caso si dice che il sistema ha infinite soluzioni e che presenta una variabile libera (la z in questo caso, ma procedendo con un ordine diverso nelle sostituzioni poteva essere la x o la y).

L'insieme delle soluzioni del sistema è quindi:

$$S = \left\{ \left(\frac{7-2z}{5}, \frac{8-3z}{5}, z\right) \mid z \in \mathbb{R} \right\} \subset \mathbb{R}^3$$

In generale, per "sistema lineare" si intende un sistema composto di equazioni lineari, cioè date da polinomi di primo grado nelle incognite.

Vedremo che esso può ammettere infinite soluzioni (sistema *indeterminato*), oppure nessuna soluzione (sistema *impossibile*) oppure un'unica soluzione (sistema *determinato*). Ricordiamo che con la parola *soluzione*, per un sistema a n incognite, si intende una n-pla di numeri reali che, assegnati come valori alle incognite, rendano vere le equazioni del sistema.

Ad esempio il seguente sistema è impossibile (verificarlo per esercizio):

$$\begin{cases} 2x - 3y = 12 \\ 4x - 6y - 3 = 0 \end{cases}.$$

I sistemi lineari sono una forma di modello matematico di varie situazioni (economiche, fisiche) e, per quanto non molto complessi, sono largamente utilizzati in pratica.

Esempio 4.1.2. Una banca ha investito 120 milioni di euro nelle azioni delle società Alcanada, Burraco, Cordoline e Delta Airlines. La somma degli investimenti nelle azioni di A e C è stata pari a quella nelle B e D. L'investimento in D è stato il doppio di quello in B. A fine anno, la banca comunica che ha guadagnato il 20% del capitale investito nelle azioni A, il 10% del capitale investito nelle azioni B, il 5% di quello investito nelle D e infine ha perso il 10% del capitale investito nelle azioni C.

Sapendo che la banca ha realizzato un guadagno totale di 7 milioni di euro, si vuole determinare la composizione del portfolio di investimento.

Nel risolvere un problema di questo tipo per prima cosa vanno individuate le incognite che vogliamo determinare; in questo caso si tratta dei valori (in milioni di euro) degli investimenti nelle società A, B, C e D, che chiameremo, rispettivamente, x_1, x_2, x_3 e x_4.

I dati del problema si traducono a questo punto nelle seguenti equazioni:

$$\begin{array}{cc} x_1 + x_2 + x_3 + x_4 = 120 & \text{[totale dell'investimento]} \\ x_1 + x_3 = x_2 + x_4 & \text{[prima relazione data fra gli investimenti]} \\ x_4 = 2x_2 & \text{[seconda relazione data fra gli investimenti]} \\ \frac{x_1}{5} + \frac{x_2}{10} - \frac{x_3}{10} + \frac{x_4}{20} = 7 & \text{[totale del profitto conseguito]} \end{array}$$

Riordinate e scritte più propriamente, queste equazioni danno il sistema:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 120 \\ x_1 - x_2 + x_3 - x_4 = 0 \\ 2x_2 - x_4 = 0 \\ 4x_1 + 2x_2 - 2x_3 + x_4 = 140 \end{cases}$$

Possiamo risolverlo per sostituzione, come abbiamo fatto nel primo esempio, ma proviamo un'altra strada: sottraiamo la prima equazione alla seconda, e alla quarta il quadruplo della prima (così elimineremo la x_1 dalle due equazioni, semplificando il sistema). Otterremo:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 120 \\ -2x_2 - 2x_4 = -120 \\ 2x_2 - x_4 = 0 \\ -2x_2 - 6x_3 - 3x_4 = -340 \end{cases}$$

Poiché dalla terza equazione si ha $x_4 = 2x_2$, dalla seconda si ottiene $6x_2 = 120$, e quindi $x_2 = 20$ e $x_4 = 40$. Sostituendo questi valori nella quarta equazione si ricava $6x_3 = 180$ e cioè $x_3 = 30$. Infine sostituendo nella prima equazione i valori trovati si ha $x_1 = 30$ e quindi la soluzione (unica) del sistema è

$$\mathcal{S} = \{(30, 20, 30, 40)\}.$$

Abbiamo qui sopra seguito un metodo "misto"; in parte costituito dal combinare le equazioni per semplificarle (togliendo variabili) e in parte dal metodo di sostituzione.

Quello che cerchiamo è un buon metodo (un algoritmo) che sia utilizzabile sempre (e magari sia implementabile in un computer).

4.2 Risoluzione di sistemi lineari

In generale, un sistema lineare di m equazioni e n incognite si scriverà così:

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{cases} \tag{4.1}$$

Ove le $x_1, \ldots, x_n$ sono le incognite, gli a_{ij} sono i loro coefficienti e i $b_1, \ldots, b_m$ sono i termini noti.

Una soluzione di (4.1) è una n-upla di $\mathbb{R}^n$, che assegnata come n-upla di valori per le x_i soddisfa tutte le equazioni di (4.1). Come abbiamo notato prima, si ha che

Proposizione 4.2.1. *Se un sistema di tipo (4.1) ammette soluzioni, esse sono una sola oppure infinite.*

Per la dimostrazione di ciò si rimanda al Teorema 4.3.1, di cui la Proposizione 4.2.1 è una conseguenza immediata.

Come si risolve un sistema di questo tipo? Il metodo più intuitivo è quello che abbiamo usato nel primo esempio analizzato; quello per *sostituzione*. Dalla prima equazione si può ricavare

$$x_1 = \frac{-1}{a_{11}}(a_{12}x_2 + \cdots + a_{1n}x_n - b_1)$$

e la si sostituisce nella seconda equazione, dove si ricava la x_2, e così via (questo se $a_{11} \neq 0$, altrimenti si ricava la prima delle x_i che abbia il coefficiente $\neq 0$).

Questa procedura è senz'altro corretta, ma un calcolatore che dovesse risolvere un sistema di centinaia di equazioni con centinaia di incognite (simili sistemi non sono inusuali in molte applicazioni pratiche) impiegherebbe anni per farlo. Quindi la procedura, benchè corretta, non è efficiente, e non è questo l'algoritmo che cerchiamo per risolvere i sistemi lineari.

Abbiamo visto nell'Esempio 4.1.2, che può essere utile combinare le equazioni che abbiamo per eliminarne delle variabili. Un algoritmo basato su questa idea e molto efficiente è l'*eliminazione di Gauss*.

Si tratta di effettuare combinazioni delle equazioni del sistema al fine di semplificarlo, riducendo il numero di incognite presenti nelle equazioni. Introduciamo questa procedura partendo da un esempio:

$$\begin{cases} x & + & y & + & z & = & 3 \\ 2x & + & 3y & - & z & = & 4 \\ x & + & 2y & + & 2z & = & 5 \end{cases}$$

utilizziamo la prima equazione per eliminare la x dalle altre due, sottraendola dalla terza e, moltiplicata per due, dalla seconda. Otterremo:

$$\begin{cases} x + y + z = 3 \\ \quad y - 3z = -2 \\ \quad y + z = 2 \end{cases}$$

Ripetendo il procedimento, si usa la seconda equazione per eliminare la y dalla terza, ottenendo:

$$\begin{cases} x + y + z = 3 \\ \quad y - 3z = -2 \\ \quad\quad 4z = 4 \end{cases}$$

Adesso nell'ultima equazione compare una sola variabile, e così possiamo risolverla ottenendo $z = 1$, che, sostituendo all'indietro, porterà alla soluzione (unica) del sistema:

$$\begin{cases} x = 1 \\ y = 1 \\ z = 1 \end{cases}.$$

Naturalmente non sempre troveremo una soluzione, oppure essa potrebbe non essere unica; per esempio, se a un certo punto del procedimento di eliminazione otteniamo una equazione del tipo $7 = 0$, allora il sistema è impossibile, cioè non ammette soluzione.

Esempio 4.2.1. Applicando l'eliminazione di Gauss al sistema

$$\begin{cases} 2x + 2y = 4 \\ x + y = 1 \end{cases}$$

si ottiene

$$\begin{cases} 2x + 2y = 4 \\ 0 = -1 \end{cases},$$

e quindi esso non ha soluzioni.

Quando invece l'algoritmo di eliminazione porta a un'ultima equazione che contiene più di una variabile, esprimeremo la prima di esse in funzione delle altre, che rimarranno libere di variare e saranno dette appunto *variabili libere* (come nell'Esempio 4.1.1); il sistema ha quindi in questo caso infinite soluzioni.

Definizione 4.2.1. Quando un sistema lineare ha infinite soluzioni, dipendenti da t variabili libere, si dice che esso ha "∞^t soluzioni".

Esempio 4.2.2.

$$\begin{cases} y+z=3 \\ x+y+z=4 \end{cases}.$$

Innanzitutto qui non possiamo usare la prima equazione per iniziare l'eliminazione, perché essa manca della x, quindi facciamo un'inversione nell'ordine delle equazioni (il che non cambia le soluzioni del sistema, ed è un artificio a cui ricorreremo sempre quando una variabile non sia presente ove dovremmo usarla per eliminarla dalle altre equazioni). Il sistema adesso sarà già in forma ridotta (la x non compare nella seconda equazione).

$$\begin{cases} x + y + z = 4 \\ \quad\; y + z = 3 \end{cases} \rightarrow \begin{cases} x + y + z = 4 \\ \quad\; y \quad\; = 3-z \end{cases}$$

e, sostituendo nella prima equazione, si ottiene:

$$\begin{cases} x = z-3-z+4=1 \\ y = 3-z \end{cases} \rightarrow \begin{cases} x=1 \\ y=3-z \end{cases}.$$

Quindi le soluzioni sono tutte le triple del tipo: $(1, 3-z, z)$, cioè il sistema ha ∞^1 soluzioni, e l'insieme S delle soluzioni si può scrivere:

$$S = \{(1, 3-z, z) \mid z \in \mathbb{R}\} \subset \mathbb{R}^3.$$

Esercizio 4.2.1. Bilanciare la seguente reazione chimica:

$$\mathrm{Cu} + \mathrm{H_2SO_4} \rightarrow \mathrm{CuSO_4} + \mathrm{H_2O} + \mathrm{SO_2}$$

Soluzione: siano $x_1, x_2, \ldots, x_5$ i coefficienti da determinare per le 5 molecole che appaiono nella reazione.

$$x_1\mathrm{Cu} + x_2\mathrm{H_2SO_4} \rightarrow x_3\mathrm{CuSO_4} + x_4\mathrm{H_2O} + x_5\mathrm{SO_2}.$$

Esaminando come i singoli elementi appaiono ai due lati della reazione, si ottengono le seguenti uguaglianze:

$$\mathrm{Cu} \rightarrow x_1 = x_3;$$

$$\mathrm{H} \rightarrow 2x_2 = 2x_4;$$

$$\mathrm{S} \rightarrow x_2 = x_3 + x_5;$$

$$\mathrm{O} \rightarrow 4x_2 = 4x_3 + x_4 + 2x_5;$$

quindi il sistema risultante sarà:

$$\begin{cases} x_1 - x_3 = 0 \\ 2x_2 - 2x_4 = 0 \\ x_2 - x_3 - x_5 = 0 \\ 4x_2 - 4x_3 - x_4 - 2x_5 = 0 \end{cases}.$$

L'eliminazione porta alla forma (scritta in modo da evidenziare la posizione delle variabili):

$$\begin{cases} x_1 - x_3 = 0 \\ x_2 - x_4 = 0 \\ -x_3 + x_4 - x_5 = 0 \\ -x_4 - 2x_5 = 0 \end{cases}.$$

Dall'ultima equazione si ha $x_4 = 2x_5$, e da qui:

$$\begin{cases} x_1 = x_5 \\ x_2 = 2x_5 \\ x_3 = x_5 \\ x_4 = 2x_5 \end{cases}.$$

Le soluzioni sono ∞^1, date da

$$\mathcal{S} = \{(x_5, 2x_5, x_5, 2x_5, x_5) \mid x_5 \in \mathbb{R}\}.$$

A noi interesseranno solo quelle intere, e in particolare la più piccola possibile; poniamo quindi la variabile libera $x_5 = 1$, e otteniamo la soluzione $(1, 2, 1, 2, 1)$, che ci dà:

$$\mathrm{Cu} + 2\mathrm{H_2SO_4} \rightarrow \mathrm{CuSO_4} + 2\mathrm{H_2O} + \mathrm{SO_2}$$

Consideriamo di nuovo il generico sistema lineare (4.1), che abbiamo considerato all'inizio della sezione, e cerchiamo di riscriverlo utilizzando una notazione matriciale. La matrice dei coefficienti delle equazioni di (4.1) è:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix},$$

mentre le matrici

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

danno, rispettivamente, le incognite e i termini noti. Adesso il sistema si può rappresentare con l'equazione matriciale (verificare):

$$A \cdot X = B \tag{4.2}$$

Ad esempio, il sistema:

$$\begin{cases} x+y+z=4 \\ x+2y-z=1 \\ 3x+3y-z=4 \end{cases}$$

si può scrivere come in (4.2):

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 3 & 3 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 4 \end{pmatrix}$$

Osserviamo che le matrici A e B contengono tutta l'informazione necessaria per individuare il sistema dato; in effetti adesso l'operazione di eliminazione di Gauss si può eseguire sulla matrice

$$A|B = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 1 & 2 & -1 & 1 \\ 3 & 3 & -1 & 4 \end{array}\right),$$

nella quale abbiamo riunito i dati salienti della (4.2). ottenendo la nuova matrice $A|B$ semplicemente accostando ad A la colonna B dei termini noti.

La riduzione si può effettuare ora sulla matrice come sulle equazioni, cioè utilizzando la prima riga per eliminare gli elementi sulla prima colonna (dal secondo in poi), e proseguendo con l'utilizzare la seconda riga per eliminare gli elementi della seconda colonna (dal terzo in poi), e così via, ottenendo la matrice:

$$A'|B' = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & -4 & -8 \end{array}\right),$$

che corrisponde al sistema:

$$\begin{cases} x+y+z=4 \\ \quad y-2z=-3 \\ \quad\quad -4z=-8 \end{cases}$$

da cui: $z=2$, e (sostituendo): $y-4=-3$, cioè $y=1$; e infine, dalla prima equazione, $x+1+2=4$, cioè: $x=1$.

Quindi la soluzione (unica) del sistema è: $X = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

Il vantaggio che si ottiene con questa formalizzazione è che potremo dire (vedi più avanti) se il sistema ha soluzioni, e quante ne ha, anche senza risolverlo, solo studiando le matrici A e B.

Chiariamo adesso che cosa intendiamo per *eliminazione di Gauss* su di una matrice $A \in \mathbb{R}^{m,n}$; essa consiste nel fare due soli tipi di operazione:

- sommare a una riga un multiplo di una riga precedente;
- scambiare fra loro due righe.

Operiamo così al fine di trasformare la nostra matrice A in una matrice triangolare superiore A', avente solo degli zero sotto la diagonale principale (cioè nelle posizioni a'_{ij} con $i > j$).

Consideriamo un esempio:

$$A = \begin{pmatrix} 1 & 1 & 0 & 3 & 5 \\ 1 & 2 & 1 & -1 & -2 \\ 2 & 3 & 1 & 2 & 3 \\ 2 & 2 & 2 & 2 & 2 \end{pmatrix}$$

per fare il primo passo dell'eliminazione di Gauss utilizziamo la prima riga per eliminare nella prima colonna tutti gli elementi diversi da zero (tranne il primo); dovremo sommare alla seconda riga la prima moltiplicata per -1, alla terza e alla quarta la prima moltiplicata per -2; otterremo:

$$\begin{pmatrix} 1 & 1 & 0 & 3 & 5 \\ 0 & 1 & 1 & -4 & -7 \\ 0 & 1 & 1 & -4 & -7 \\ 0 & 0 & 2 & -4 & -8 \end{pmatrix}$$

ora usiamo la seconda riga per eliminare nella terza e quarta il secondo elemento (nella quarta esso è già zero, quindi non abbiamo da fare niente):

$$\begin{pmatrix} 1 & 1 & 0 & 3 & 5 \\ 0 & 1 & 1 & -4 & -7 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -4 & -8 \end{pmatrix}$$

quello che dovremmo fare adesso è utilizzare la terza riga per eliminare il terzo elemento ($= 2$) dalla quarta, ma questo è impossibile perché la terza riga è fatta di soli zero; allora scambiamo di posto la terza e la quarta riga:

$$\begin{pmatrix} 1 & 1 & 0 & 3 & 5 \\ 0 & 1 & 1 & -4 & -7 \\ 0 & 0 & 2 & -4 & -8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = A'$$

e abbiamo ottenuto la matrice A' della forma voluta.

Consideriamo un altro esempio:

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 1 & 1 & -2 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & -1 & -1 \end{pmatrix}$$

in questo caso le prime due colonne sono nulle, quindi dovremo cominciare a eliminare sulla terza, ma è impossibile farlo utilizzando la prima riga, in quanto in essa anche il terzo elemento è 0; scambieremo quindi la prima riga con la seconda, e procederemo con l'eliminazione sulla terza colonna:

$$\begin{pmatrix} \mathbf{0} & 0 & 1 & 1 & 1 & -2 & -1 \\ \mathbf{0} & 0 & 0 & 1 & 0 & 2 & 2 \\ \mathbf{0} & 0 & 1 & 0 & 0 & 0 & 0 \\ \mathbf{0} & 0 & 2 & 0 & 1 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 0 & \mathbf{0} & 1 & 1 & 1 & -2 & -1 \\ 0 & \mathbf{0} & 0 & 1 & 0 & 2 & 2 \\ 0 & \mathbf{0} & 0 & -1 & -1 & 2 & 1 \\ 0 & \mathbf{0} & 0 & -2 & -1 & 3 & 1 \end{pmatrix}$$

proseguiamo l'eliminazione sulla quarta colonna (utilizzando la seconda riga):

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 & -2 & -1 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 2 & 2 \\ 0 & 0 & 0 & -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & -2 & -1 & 3 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & -2 & -1 \\ 0 & 0 & 0 & 1 & \mathbf{0} & 2 & 2 \\ 0 & 0 & 0 & 0 & -1 & 4 & 3 \\ 0 & 0 & 0 & 0 & -1 & 7 & 5 \end{pmatrix}$$

terminiamo eliminando sull'ultima riga in quinta posizione (utilizzando la penultima riga):

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 & -2 & -1 \\ 0 & 0 & 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & -1 & 4 & 3 \\ 0 & 0 & 0 & 0 & -1 & 7 & 5 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & -2 & -1 \\ 0 & 0 & 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & -1 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 & 3 & 2 \end{pmatrix}.$$

4.3 Algoritmo per l'eliminazione di Gauss

Diamo adesso in modo dettagliato i passi che costituiscono l'algoritmo che descrive come eseguire l'eliminazione di Gauss su una matrice A:

i) Consideriamo la prima colonna non nulla di A; sia essa di indice j_1. Avremo due possibilità: o in tale colonna il primo elemento (cioè a_{1j_1}) è diverso da zero, oppure no, e sia $a_{i_1 j_1}$ il primo elemento non nullo (con $i_1 > 1$). Nel primo caso passiamo al passo iii), nel secondo caso al passo ii).

ii) Scambiamo fra loro la prima e la i_1-esima riga, ottenendo una nuova matrice ove l'elemento di posto $1j_1$ è diverso da zero.

iii) Eliminiamo nella colonna j_1-esima tutti gli elementi diversi da zero dalla seconda riga in poi, lasciando inalterata la prima riga, e usandola per fare l'eliminazione sulle altre, ottenendo una nuova matrice con $a_{ij_1} = 0$ per $i = 2, \ldots, m$.

iv) Lasciando ora inalterata la prima riga, ripetiamo i passi i), ii), iii) sulla matrice $(m-1) \times n$ delle altre $m-1$ righe, a meno che esse non siano tutte nulle, nel qual caso ci fermiamo.

Consideriamo a cosa arriviamo con questo procedimento; da una matrice $A \in \mathbb{R}^{m,n}$, esso darà una matrice $A' = (a'_{ij}) \in \mathbb{R}^{m,n}$, che sarà sicuramente triangolare superiore (ha solo degli "0" sotto la diagonale, cioè nelle posizioni a'_{ij} con $i > j$); ma l'effetto della riduzione di Gauss dà anche qualcosa di altro: porterà la matrice in forma di *matrice a gradini*, cioè del tipo:

$$\begin{pmatrix}
0 & \dots & \dots & 0 & \underline{|p_1} * & \dots & \dots & \dots & \dots & \dots & \dots & * \\
0 & \dots & \dots & 0 & \dots & 0 & \underline{|p_2} * & \dots & \dots & \dots & \dots & * \\
0 & \dots & \dots & 0 & \dots & 0 & \dots & 0 & \underline{|p_3} * & \dots & \dots & * \\
 & \vdots & & & & & \vdots & & & \vdots & & \\
0 & \dots & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \underline{|p_r} * & * \\
0 & \dots & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\
 & \vdots & & & & & \vdots & & & \vdots & & \\
0 & \dots & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0
\end{pmatrix}$$

ove, nelle notazioni di iii), $p_1 = a'_{1,j_1}$.

La nostra matrice non è solo triangolare superiore, ma ha anche la proprietà che sotto al primo elemento non nullo di ogni riga compaiono solo zeri (questo si intende con *matrice a gradini*).

Abbiamo indicato con $p_1, p_2, \dots, p_r$ i primi elementi diversi da zero in ogni riga di una matrice a gradini; essi prendono il nome di *pivot*.

Notiamo che ovviamente la nozione di *pivot* ha senso solo nelle matrici a gradini, e che se r è il numero di pivot (cioè il numero di righe non nulle), allora $r \leq m$, e che se $r = m$, allora non compaiono righe nulle finali (come quelle che abbiamo indicato nella matrice sopra).

Avremo anche $r \leq n$, poiché i p_i stanno uno per colonna, quindi non possono essere più delle colonne di A.

Definizione 4.3.1. Sia $A \in \mathbb{R}^{m,n}$, e A' la matrice ottenuta dalla riduzione di Gauss di A. Il numero di pivot che appaiono nella matrice A' si dice *rango* della matrice A, e si indica con $r = r(A)$.

Il rango di una matrice è un concetto molto importante; in un certo senso si tratta di una misura della quantità di informazione portata

dalla matrice. Ad esempio, quando si tratta della matrice di un sistema, il fatto che con l'eliminazione di Gauss rimangano solo r righe non nulle ci dice che in realtà quel sistema ha solo r equazioni significative, non m: le altre erano combinazioni di quelle r, e con opportune operazioni si sono annullate.

Osservazione 4.3.1. Osserviamo che avremo $r(A) = 0$ se e solo se A' è la matrice nulla.

Vediamo adesso in che modo sia possibile, come annunciato, determinare se un sistema abbia soluzioni e quante esse siano, analizzando la matrice $(A|B)$.

Esempio 4.3.1. Consideriamo il sistema:

$$\begin{cases} x - y + z = 3 \\ x + y - 2z = 4 \\ 3x - y = 5 \end{cases}$$

applicando Gauss otteniamo:

$$\begin{cases} x - y + z = 3 \\ 2y - 3z = 1 \\ 2y - 3z = -4 \end{cases} \to \begin{cases} x - y + z = 3 \\ 2y - 3z = 1 \\ 0 = -5 \end{cases}$$

quindi il sistema è impossibile. Vediamo la stessa cosa operando sulla matrice $(A|B)$:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 3 \\ 1 & 1 & -2 & 4 \\ 3 & -1 & 0 & 5 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -1 & 1 & 3 \\ 0 & 2 & -3 & 1 \\ 0 & 2 & -3 & -4 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -1 & 1 & 3 \\ 0 & 2 & -3 & 1 \\ 0 & 0 & 0 & -5 \end{array}\right).$$

Quindi se il sistema è impossibile, l'eliminazione di Gauss porta a un'equazione (eliminando le variabili) del tipo $0 = -5$, questo vale a dire che eseguendo l'eliminazione di $(A|B)$ si trova una riga $(0\ 0\ \dots\ 0\ 0\ p)$, con $p \neq 0$, e cioè un pivot p sull'ultima colonna, come nell'esempio precedente ove i pivot non nulli sono $1, 2, -5$ e il -5 sta nell'ultima colonna.

Si ha quindi:

IL SISTEMA NON HA SOLUZIONI
se e solo se
nella riduzione di Gauss di $(A|B)$ c'è un pivot sull'ultima colonna.

Che cosa vuol dire ciò in termini di rango delle matrici in questione? Se $(A|B)$ ha rango r, allora A avrà rango $r - 1$, perché degli r pivot di $(A'|B')$, l'ultimo sta su B', e quindi A ha $r - 1$ pivot. Allora possiamo riesprimere l'affermazione precedente così:

IL SISTEMA NON HA SOLUZIONI
se e solo se
$r(A) \neq r(A|B)$.

E se $r(A) = r(A|B) = r$? Allora i pivot di $(A'|B')$ stanno tutti nella parte A', e in questo caso l'eliminazione di Gauss porta $(A'|B')$ in forma:

$$\left(\begin{array}{ccccccccccc|c} 0 & \dots & 0 & \underline{p_1} & * & \dots & \dots & \dots & \dots & \dots & * & * \\ 0 & \dots & 0 & \dots & 0 & \underline{p_2} & * & \dots & \dots & \dots & * & * \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \underline{p_3} & * & \dots & * & * \\ & \vdots & & & \vdots & & & & \vdots & & & \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \underline{p_r} & * & * \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \\ & \vdots & & & \vdots & & & & \vdots & & & \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 \end{array}\right).$$

Ognuno dei pivot $p_1, \dots, p_r$ è il coefficiente di un'incognita e permette di ricavare quell'incognita in funzione delle altre (nell'equazione corrispondente a quella riga), quindi in tutto rimarranno $n - r$ variabili libere, cioè il sistema ha ∞^{n-r} soluzioni (notare che se $n = r$, ciò significa che abbiamo zero variabili libere, cioè che il sistema ha un'unica soluzione).

Riassumendo, abbiamo dimostrato il seguente teorema, noto come *teorema di Rouché–Capelli*

Teorema 4.3.1 (Teorema di Rouché–Capelli). *Il sistema lineare $A \cdot X = B$ ha soluzioni se e solo se*

$$r(A) = r(A|B).$$

In questo caso le soluzioni sono ∞^{n-r} (cioè ci sono $n-r$ variabili libere, e c'è una sola soluzione se $n = r$).

Esempio 4.3.2. Dire per quali valori del parametro k il seguente sistema ha soluzioni e quante ne ha:

$$\begin{cases} x + y = 2 \\ x + (k+1)y + z = 3 \\ 2x + (k+2)y + z = k + 3 \end{cases}.$$

Eseguiamo l'eliminazione di Gauss sulla matrice $(A|B)$:

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 1 & k+1 & 1 & 3 \\ 2 & k+2 & 1 & k+3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & k & 1 & 1 \\ 0 & k & 1 & k+1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & k & 1 & 1 \\ 0 & 0 & 0 & k \end{array}\right)$$

si ha che se $k \neq 0$ l'ultimo pivot sta sull'ultima colonna, e cioè che $r(A) \neq r(A|B)$, quindi il sistema non ha soluzioni. Se invece $k = 0$, allora $r(A) = r(A|B) = 2$, e quindi il sistema ha ∞^1 soluzioni.

In particolare, per $k = 0$, le soluzioni sono date da

$$\begin{cases} x + y = 2 \\ z = 1 \end{cases} \rightarrow \begin{cases} x = 2 - y \\ z = 1 \end{cases}$$

cioè le soluzioni del sistema sono le infinite terne: $\{(2 - y, y, 1)\}$ che si ottengono dando a y un qualsiasi valore in $\mathbb{R}$.

4.4 Sistemi omogenei

Considereremo qui un tipo particolare di sistemi lineari: quelli in cui tutti i termini noti sono nulli, che costituiscono un importante strumento nello studio degli spazi vettoriali.

Definizione 4.4.1. Un sistema lineare del tipo $A \cdot X = B$, con $A \in \mathbb{R}^{m,n}$, $B \in \mathbb{R}^{m,1}$, $X \in \mathbb{R}^{n,1}$ si dice *omogeneo* se $B = O$.

Esempio 4.4.1. Consideriamo il sistema omogeneo:

$$\begin{cases} x - y + z = 0 \\ 2x - y - 2z = 0 \end{cases} \rightarrow \begin{pmatrix} 1 & -1 & 1 \\ 2 & -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

applicando Gauss a A si ottiene il sistema:

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -4 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \begin{cases} x - y + z = 0 \\ y - 4z = 0 \end{cases},$$

da cui:

$$\begin{cases} x = 4z - 2z = 3z \\ y = 4z \\ z = z \end{cases}$$

quindi abbiamo una variabile libera, z, e le soluzioni del sistema sono date dall'insieme $V \subset \mathbb{R}^n$:

$$V = \{(x, y, z) \mid y = 4z,\ x = 3z\} = \{(3z, 4z, z) \in \mathbb{R}^3 \mid z \in \mathbb{R}\}.$$

È facile verificare che V è un sottospazio di $\mathbb{R}^3$ (farlo per esercizio); una base di V è data da: $\{(3, 4, 1)\}$, poiché ogni $\mathbf{v} \in V$ è del tipo $(3z, 4z, z)$ e si ottiene come $z(3, 4, 1)$.

I fatti osservati nel precedente esempio sono del tutto generali, come afferma il seguente risultato:

Proposizione 4.4.1. *Sia $A \cdot X = O$ un sistema lineare omogeneo (come nella Definizione 4.4.1); allora l'insieme V delle soluzioni del sistema è un sottospazio vettoriale di $\mathbb{R}^n$, e la sua dimensione è pari al numero delle variabili libere del sistema, e cioè $n - r(A)$.*

Dimostrazione. È banale osservare che un sistema omogeneo ammette sempre la soluzione nulla, quindi $O \in V \subset \mathbb{R}^n$.
Siano ora $X_1, X_2 \in V$ due soluzioni del sistema; allora

$A \cdot (X_1 + X_2) = A \cdot X_1 + A \cdot X_2 = O + O = O$, e,

$\forall \alpha \in \mathbb{R}$, $A \cdot (\alpha X_1) = \alpha A \cdot X_1 = \alpha O = O$;

quindi $X_1 + X_2 \in V$ e $\alpha X_1 \in V$. Allora V è un sottospazio vettoriale di $\mathbb{R}^n$ per la Proposizione 2.4.1.

Se risolviamo con l'eliminazione di Gauss il sistema, sappiamo (dal Teorema di Rouché–Capelli) che troveremo $n - r(A)$ variabili libere (nel caso dei sistemi omogenei abbiamo sempre soluzioni del sistema poiché $A \cdot O_n = O_m$, quindi $O_n \in \mathbb{R}^n$ è sempre soluzione). Se $n = r(A)$, allora l'unica soluzione è $X = O$, cioè $V = \{O\}$, e $\dim V = 0 = n - r$, come richiesto.

Se $n > r(A)$, allora le $\infty^{n-r(A)}$ soluzioni si scriveranno in funzione delle $n - r(A)$ variabili libere $x_{r+1}, \ldots, x_n$. Per ottenere una base di V si possono assegnare i valori $\{1, 0, \ldots, 0\}$; $\{0, 1, \ldots, 0\}$; $\ldots$; $\{0, \ldots, 0, 1\}$ alle $n - r$ variabili libere ottenendo $n - r$ vettori che generano tutti gli altri di V e sono fra loro indipendenti, quindi $\dim V = n - r(A)$. □

Esempio 4.4.2. Sia $V = \{(x, y, z, t) \mid x - 2y + z = 0; y + z + t = 0\} \subset \mathbb{R}^4$. Risolvendo il sistema

$$\begin{cases} x - 2y + z = 0 \\ y + z + t = 0 \end{cases}$$

troviamo $y = -z - t$, $x = -3z - 2t$, quindi due variabili libere:

$$V = \{(-3z - 2t, -z - t, z, t)\}.$$

Assegnando a z e a t i valori $1, 0$ e $0, 1$ (rispettivamente) otteniamo i vettori $\mathbf{v}_1 = (-3, -1, 1, 0)$ e $\mathbf{v}_2 = (-2, -1, 0, 1)$. Essi sono una base di V in quanto sono indipendenti e ogni vettore di V si scrive come: $z\mathbf{v}_1 + t\mathbf{v}_2$.

4.5 Rango di una matrice

Riconsidereremo in questa sezione i teoremi che descrivono le soluzioni dei sistemi lineari di equazioni in termini di concetti della teoria degli spazi vettoriali.

L'importante quantità $r(A)$, il rango di una matrice A, si può definire anche in un modo più intrinsecamente legato alla matrice (senza riferirci all'eliminazione di Gauss):

Proposizione 4.5.1. *Il rango di una matrice $A \in \mathbb{R}^{m,n}$ è uguale al massimo numero di righe di A linearmente indipendenti (viste come vettori di $\mathbb{R}^n$) o, equivalentemente, al massimo numero di colonne di A linearmente indipendenti (viste come vettori di $\mathbb{R}^m$). In sintesi:*

$$\dim \mathcal{R}(A) = \dim \mathcal{C}(A) = r(A).$$

Dimostrazione. Come abbiamo già visto nel Capitolo 3, a ogni matrice $A \in \mathbb{R}^{m,n}$ si possono associare due sottospazi: $\mathcal{R}(A) \subset \mathbb{R}^n$, lo spazio generato dalle righe di A, e $\mathcal{C}(A) \subset \mathbb{R}^m$, lo spazio generato dalle colonne di A. La dimostrazione della proposizione si articola in tre passi:

1. Sia A' ottenuta da A tramite l'eliminazione di Gauss, allora:
$$\mathcal{R}(A) = \mathcal{R}(A').$$
2. $$\dim \mathcal{R}(A) = r(A),$$
3. $$\dim \mathcal{C}(A) = \dim \mathcal{R}(A).$$

Dimostriamo il punto 1. Se consideriamo i vettori riga di A e quelli di A', si ha che essi generano lo stesso sottospazio di $\mathbb{R}^n$, in quanto ogni passo dell'eliminazione di Gauss consiste o nello scambiare fra loro due righe (e questo ovviamente non altera lo spazio generato dalle righe stesse), oppure nel sommare a una riga un multiplo di un'altra, cioè se le righe di A sono i vettori $\mathbf{R_1}, \ldots, \mathbf{R_m}$, si avrà che si passa allo spazio $\langle \mathbf{R_1}, \ldots, \mathbf{R_i} + \lambda \mathbf{R_j}, \ldots, \mathbf{R_m} \rangle$, ove alla i-esima riga si è aggiunta la j-esima ($i \neq j$) moltiplicata per $\lambda \neq 0$. Ovviamente con le $\mathbf{R_1}, \ldots, \mathbf{R_m}$ si possono generare le $\mathbf{R_1}, \ldots, \mathbf{R_i} + \lambda \mathbf{R_j}, \ldots, \mathbf{R_m}$ (l'unico vettore diverso è l'i-esimo che è combinazione lineare di $\mathbf{R_i}$ e $\mathbf{R_j}$); ma vale anche il viceversa: con $\mathbf{R_j}$ e $\mathbf{R_i} + \lambda \mathbf{R_j}$ si può generare $\mathbf{R_i} = (\mathbf{R_i} + \lambda \mathbf{R_j}) - \lambda \mathbf{R_j}$.

Allora lo spazio generato dalle righe non è cambiato e il numero di righe indipendenti di A è lo stesso di quelle di A' e quindi il punto 1 è dimostrato.

Vediamo ora il punto 2. Sappiamo che $r(A) = r(A')$ per definizione e che $\mathcal{R}(A) = \mathcal{R}(A')$ dal punto 1), quindi possiamo lavorare su A'.

Quante sono le righe indipendenti di A'? In A' abbiamo $r = r(A')$ righe non nulle (ciascuna con un pivot p_i, $i = 1, \ldots, r$) e $m - r$ righe nulle:

$$A' = \begin{pmatrix}
0 & \ldots & 0 & \underline{|p_1|} * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
0 & \ldots & 0 & \ldots & 0 & \underline{|p_2|} * & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \underline{|p_3|} * & \ldots & \ldots & \ldots & * \\
 & \vdots & & & & \vdots & & & & \vdots & & \vdots \\
0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \underline{|p_r|} * & \ldots & * \\
0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & \ldots & 0 \\
 & \vdots & & & & \vdots & & & & \vdots & & \vdots \\
0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & \ldots & 0
\end{pmatrix}.$$

Ovviamente possiamo trascurare le righe nulle. Vediamo che le r righe non nulle sono indipendenti. Siano $\mathbf{R}_1, \ldots, \mathbf{R}_r$ tali righe, e sia

$$\alpha_1 \mathbf{R}_1 + \cdots + \alpha_r \mathbf{R}_r = (0, \ldots, 0) \in \mathbb{R}^n$$

una loro combinazione lineare nulla. In tale combinazione dovremo avere $\alpha_1 = 0$, poiché nella posizione di p_1 (in $\mathbf{R}_1$), $\alpha_1 p_1$ è l'unico elemento non nullo (le altre righe hanno tutte uno zero sotto a p_1), allora la combinazione è:

$$\alpha_2 \mathbf{R}_2 + \cdots + \alpha_r \mathbf{R}_r = (0, \ldots, 0),$$

ma adesso possiamo ripetere il ragionamento di prima per avere $\alpha_2 = 0$, osservando p_2, e così via per tutti gli α_i, ottenendo che la combinazione deve avere tutti i coefficienti nulli, e quindi che le righe $\mathbf{R}_1, \ldots, \mathbf{R}_r$ sono linearmente indipendenti, e quindi una base per $\mathcal{R}(A')$, cosicché anche il punto 2 è dimostrato.

La dimostrazione del punto 3, e cioè che il rango è anche $\dim \mathcal{C}(A)$ (il numero di colonne linearmente indipendenti) è leggermente più complessa. Notiamo che lo spazio delle righe di A coincide con lo spazio delle colonne della trasposta ${}^tA$ e viceversa, e che facendo la trasposta di un prodotto di matrici si ha: ${}^t(A \cdot B) = {}^tB \cdot {}^tA$.

Supponiamo ora che $r(A) = \dim(\mathcal{R}(A)) = k$, e sia $\{\mathbf{S}_1, \ldots, \mathbf{S}_k\}$ una base di $\mathcal{R}(A)$; allora ogni riga $\mathbf{R}_i$ di A sarà combinazione lineare di $\mathbf{S}_1, \ldots, \mathbf{S}_k$, cioè esisteranno dei $b_{ij} \in \mathrm{R}$ tali che:

$$\forall i \in \{1, \ldots m\}, \quad \mathbf{R}_i = \sum_{j=1}^{k} b_{ij} \mathbf{S}_j.$$

questo ci dà che, detta $B = (b_{ij}) \in \mathbb{R}^{m,k}$, e $S \in \mathbb{R}^{k,n}$ la matrice che ha $\mathbf{S}_1, \ldots, \mathbf{S}_k$ per righe:

$$A = B \cdot S,$$

e allora, trasponendo, avremo che

$${}^tA = {}^tS \cdot {}^tB,$$

dalla quale si ricava che le righe di ${}^tA$ sono combinazioni lineari delle righe di ${}^tB$, che sono k, quindi $\dim \mathcal{R}({}^tA) \leq k$, cioè che

$\dim \mathcal{C}(A) = \dim \mathcal{R}({}^tA) \leq k$.

Abbiamo quindi dimostrato che, per ogni matrice A si ha:

$\dim \mathcal{C}(A) \leq \dim \mathcal{R}(A)$.

Eseguendo lo stesso ragionamento a partire da ${}^tA$ invece che da A, si ottiene anche la diseguaglianza:

$\dim \mathcal{C}({}^tA) \leq \dim \mathcal{R}({}^tA)$,

che equivale a

$\dim \mathcal{R}(A) \leq \dim \mathcal{C}(A)$,

quindi $\dim \mathcal{R}(A) = \dim \mathcal{C}(A)$, e anche il punto 3 è dimostrato.

□

Osservazione 4.5.1. Quanto visto fin qui ci dice che se abbiamo dei vettori $\mathbf{v_1}, \ldots, \mathbf{v_s} \in \mathbb{R}^n$, per vedere quanti di essi sono linearmente indipendenti (e quindi anche per vedere quanto vale $\dim\langle \mathbf{v_1}, \ldots, \mathbf{v_s}\rangle$), basta costruire una matrice A che abbia tali vettori come righe e poi eseguire l'eliminazione di Gauss su tale matrice. Il rango (e cioè il numero delle righe non nulle) della matrice A' così ottenuta darà la quantità richiesta.

Esempio 4.5.1. Determinare la dimensione di $\langle \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\rangle$ al variare del parametro $k \in \mathbb{R}$, essendo:

$$\mathbf{v}_1 = (1, k, 2k),\ \mathbf{v}_2 = (k, 2k, k+1),\ \mathbf{v}_3 = (2k, 4k, 2k).$$

Dobbiamo determinare se $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ siano o no indipendenti, perciò applichiamo Gauss alla matrice che li vede come righe:

$$\begin{pmatrix} 1 & k & 2k \\ k & 2k & k+1 \\ 2k & 4k & 2k \end{pmatrix} \to \begin{pmatrix} 1 & k & 2k \\ 0 & 2k-k^2 & k+1-2k^2 \\ 0 & 4k-2k^2 & 2k-4k^2 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & k & 2k \\ 0 & 2k-k^2 & k+1-2k^2 \\ 0 & \mathbf{0} & -2 \end{pmatrix}.$$

da qui si vede immediatamente che se $2k - k^2 \neq 0$, si hanno 3 pivot non nulli, e quindi il rango della matrice risulta essere 3, cosicchè si avrà $\dim\langle \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\rangle = 3$.

Se invece $2k - k^2 = 0$, vale a dire se $k = 0$ o $k = 2$, i tre vettori non sono indipendenti. Dobbiamo quindi discutere i due casi separatamente.

Se $k = 0$: la matrice diviene

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -2 \end{pmatrix} \text{ che dopo aver applicato Gauss diventa } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

per cui si ha $r(A) = \dim\langle \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\rangle = 2$.

Se invece $k = 2$, la matrice diviene:

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 0 & -5 \\ 0 & 0 & -2 \end{pmatrix} \text{ che dopo aver applicato Gauss diventa } \begin{pmatrix} 1 & 2 & 4 \\ 0 & 0 & -5 \\ 0 & 0 & 0 \end{pmatrix}$$

e si ha di nuovo $r(A) = \dim\langle \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\rangle = 2$.

Quanto era stato dimostrato nel teorema di Rouché–Capelli, utilizzando l'eliminazione di Gauss, può ora esser rivisto in termini di spazi vettoriali:

Un sistema lineare $AX = B$ ha soluzioni se e solo se si ha:

$$r(A) = r(A|B).$$

Dimostrazione. Siano:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix},$$

mentre

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

allora il sistema $AX = B$ si può scrivere (per colonne):

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \tag{4.3}$$

Quindi esistono dei valori di $x_1, \ldots, x_n$ che sono soluzioni del sistema, se e solo se tali valori realizzano (in (4.3)) la colonna B come combinazione lineare delle colonne di A. Ma dire che B è combinazione lineare delle colonne di A, equivale a dire che $B \in \mathcal{C}(A) \subset \mathbb{R}^m$, e cioè che aggiungendo B allo spazio $\mathcal{C}(A)$, lo spazio generato dalle colonne di A, non aumenteremo la dimensione di tale spazio, e cioè che $\mathrm{r}(A) = \mathrm{r}(A|B)$, come richiesto. □

4.6 Esercizi

Esercizio 4.6.1. Risolvere il seguente sistema lineare:

$$\begin{cases} 3x + 12y - 1988z = 9 \\ 18x + 4y + 2867z = -14 \\ 10x - 9y + 1212z = -19 \end{cases}.$$

Esercizio 4.6.2. Risolvere il seguente sistema lineare:

$$\begin{cases} x + y + z = 4 \\ x + y - z = 0 \\ 3x + 3y - z = 4 \end{cases}.$$

Esercizio 4.6.3. Risolvere il seguente sistema lineare:

$$\begin{cases} 2x + 2y - z = 3z \\ x - 3y + 2z = 0 \end{cases}.$$

Esercizio 4.6.4. Dire se il sistema lineare $A \cdot X = B$ ha o no soluzioni e quante, ove:

$$A = \begin{pmatrix} 3 & 3 & 1 \\ 2 & -1 & 1 \\ 1 & 4 & 0 \end{pmatrix}; \; B = \begin{pmatrix} 15 \\ 3 \\ 9 \end{pmatrix}; X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Esercizio 4.6.5. Dire se il sistema lineare $A \cdot X = B$ ha o no soluzioni e quante, ove:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \\ 1 & -1 & -1 \\ 2 & -1 & -3 \end{pmatrix}; \; B = \begin{pmatrix} 4 \\ 7 \\ -2 \\ -3 \end{pmatrix}; X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Esercizio 4.6.6. Discutere al variare di $k \in \mathbb{R}$ l'insieme delle soluzioni del sistema lineare $A \cdot X = B$, con:

$$A = \begin{pmatrix} 1 & 2 & k \\ 1 & k+3 & 0 \\ 1 & 2 & k \end{pmatrix}; \; B = \begin{pmatrix} k+1 \\ 3 \\ k+1 \end{pmatrix}; X = \begin{pmatrix} x \\ y \\ z \end{pmatrix};$$

Esercizio 4.6.7. Discutere al variare di $k \in \mathbb{R}$ l'insieme delle soluzioni del sistema lineare $A \cdot X = B$, con:

$$A = \begin{pmatrix} 0 & k & k \\ k & 2 & 0 \\ -3 & 1 & 0 \end{pmatrix}; \; B = \begin{pmatrix} k \\ 1 \\ -k \end{pmatrix}; X = \begin{pmatrix} x \\ y \\ z \end{pmatrix};$$

Esercizio 4.6.8. Risolvere, se possibile, il seguente sistema lineare in $\mathbb{R}^3$:

$$\begin{cases} x + z = 2 \\ 2y + z = 1 \\ x - y = 1 \\ 3x - y + z = 4 \end{cases}.$$

Esercizio 4.6.9. Risolvere, se possibile, il seguente sistema lineare in $\mathbb{R}^3$:

$$\begin{cases} x + 2y - z = 0 \\ 2x - y + z = 3 \\ x + 3y + z = 2 \end{cases}.$$

Esercizio 4.6.10. Discutere al variare del parametro $h \in \mathrm{R}$ e risolvere, ove possibile, il seguente sistema lineare in R^2:

$$\begin{cases} hx + (2h-1)y = 2 \\ x + hy = 2h \end{cases}.$$

Esercizio 4.6.11. Discutere al variare del parametro $h \in \mathrm{R}$ e risolvere, ove possibile, il seguente sistema lineare in R^3:

$$\begin{cases} x + hy + z = 2h - 1 \\ x + y + hz = 0 \\ hx + y + z = 5 \end{cases}.$$

Esercizio 4.6.12. Discutere al variare del parametro $h \in \mathbb{R}$ e risolvere, ove possibile, il seguente sistema lineare in $\mathbb{R}^2$:

$$\begin{cases} x + y = 2 \\ hx - y = -1 \\ 4x + y = 1 + h \end{cases}.$$

Esercizio 4.6.13. Discutere al variare del parametro $h \in \mathbb{R}$ e risolvere, ove possibile, il seguente sistema lineare in $\mathbb{R}^3$:

$$\begin{cases} x + (h+1)y + z = 2 + h \\ (h+3)y + z = 2h^2 \\ x + (5+h)y + (1+h)z = 6 - h \end{cases}.$$

Esercizio 4.6.14. Risolvere il sistema $A \cdot X = B$, dove:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$

Esercizio 4.6.15. Risolvere il sistema $A \cdot X = B$, dove:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 2 \\ 3 & 3 & 3 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

Esercizio 4.6.16. Discutere il seguente sistema lineare al variare di $k \in \mathbb{R}$:

$$\begin{cases} x + 2y - z - t = 0 \\ kx - y - 2kz + 2t = 1 \\ 3kx - 6kz + 3t = 2 \end{cases}.$$

Esercizio 4.6.17. Discutere il seguente sistema lineare al variare di $k \in \mathbb{R}$:

$$\begin{cases} x_1 + x_2 - kx_3 + 4x_4 = 5 \\ kx_1 + kx_2 - 3x_3 + x_4 = 3 \\ 3x_1 + 3x_2 + 5x_3 + 3x_4 = -k \end{cases}.$$

Esercizio 4.6.18. Discutere il seguente sistema lineare al variare di $h \in \mathbb{R}$:

$$\begin{cases} -x + y = -1 \\ hx + 2y = 2 \\ x + y = 3 \end{cases}$$

dove h è un parametro reale.

Esercizio 4.6.19. Discutere e, ove possibile, risolvere il seguente sistemaal variare di $h \in \mathbb{R}$:

$$\begin{cases} 2x + y - 5z = h \\ -x + 3y + hz = 0 \\ 3x + 2y - 9z = 0 \\ -x + y - 2z = 0 \end{cases}$$

dove h è un parametro reale.

Esercizio 4.6.20. Risolvere, se possibile, il seguente sistema nelle incognite x, y, z:

$$\begin{cases} x = 0 \\ y = 0 \\ 3x + 2z = 2z \end{cases}.$$

Esercizio 4.6.21. Discutere e, ove possibile, risolvere al variare di $h \in \mathbb{R}$, il seguente sistema nelle incognite x, y, z:

$$\begin{cases} x = 0 \\ y = 0 \\ hz = 0 \end{cases},$$

dove h è un parametro reale.

Esercizio 4.6.22. Sia data la matrice $A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Trovare una base dello spazio delle righe ed una base dello spazio delle colonne.

Esercizio 4.6.23. Si consideri $S \subseteq \mathbb{R}^3$, ove:

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid x + 3y - 3z = -7, 3x + y = 0\}.$$

Descrivere il generico elemento di S e dedurre che S non è un sottospazio di V.

Esercizio 4.6.24. Trovare una base del seguente sottospazio $S \subseteq \mathbb{R}^3$:

$$S := \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y = 0,\ x - z = 0,\ 2x + 2y - z = 0\}.$$

Esercizio 4.6.25. Siano V_1 e V_2 i sottospazi vettoriali di $\mathbb{R}^4$ dati dalle soluzioni dai seguenti sistemi omogenei:

$$\begin{cases} x - y + z - t = 0 \\ x + y + 2z + t = 0 \end{cases} ; \quad \begin{cases} 2x - y - z - t = 0 \\ 4x - y + 2z - t = 0 \end{cases}.$$

Determinare la dimensione e una base di $V_1 \cap V_2$.

Esercizio 4.6.26. Siano $S_1 := \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y - t = 0\}$ e $S_2 := \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y = 0,\ y - 3z = 0\}$ due sottospazi di $\mathbb{R}^4$. Trovare una base per $S_1 \cap S_2 \subseteq \mathbb{R}^4$.

Esercizio 4.6.27. Sia $h \in \mathbb{R}$ un parametro reale e siano S_1, S_2 sottospazi di $\mathbb{R}^4$ così definiti:

$$S_1 := \{(x, y, z, t) \in V \mid hy + t = 0,\ 2x - z = 0\},$$

$$S_2 := \{(x, y, z, t) \in V \mid y - hz = 0,\ (h + 1)x - 2z - ht = 0\}.$$

Stabilire la dimensione di $S_1 \cap S_2 \subseteq \mathbb{R}^4$ al variare di $h \in \mathbb{R}$.

Esercizio 4.6.28. Sia $h \in \mathbb{R}$ un parametro reale e sia $S \subset \mathbb{R}^4$ il seguente sottospazio:

$$S := \{(x, y, z, t) \in \mathbb{R}^4 \mid x - hy + ht, -2y - z + t = 0, x + y + hz = 0\}.$$

Stabilire la dimensione di S al variare di $h \in \mathbb{R}$ e trovarne una base.

Esercizio 4.6.29. Siano dati $S_1 := \{(x, y, z) \in \mathbb{R}^3 \mid x+y = x-y-z = 0\}$ e $S_2 := \{(x, y, z) \in \mathbb{R}^3 \mid 2x + kz = 0\}$ due sottoinsiemi di $\mathbb{R}^3$.

1. Mostrare che per ogni $k \in \mathbb{R}$, sia S_1 che S_2 sono sottospazi di $\mathbb{R}^3$.
2. Trovare una base per S_1 e una base per S_2.
3. Descrivere al variare di $k \in \mathbb{R}$ le dimensioni di $S_1 + S_2$ e $S_1 \cap S_2$.

Esercizio 4.6.30. * Scrivere un sistema lineare di 3 equazioni in due incognite avente $(3, 0)$ come unica soluzione. È vero che una delle equazioni è combinazione lineare delle altre? È vero che sono tutte proporzionali?

Esercizio 4.6.31. * Costruire, se esistono, quattro sistemi lineari di 2 equazioni in 3 incognite in modo che in ciascuno di essi sia verificata una delle seguenti proprietà:

1. il sistema ammetta una e una sola soluzione;
2. esistano 10 diverse soluzioni del sistema;
3. il sistema ammetta esattamente 4 soluzioni;
4. il sistema non ammetta soluzioni.

Esercizio 4.6.32. * Sia $AX = B$ un sistema lineare di m equazioni in n incognite. Stabilire quali delle seguenti affermazioni sono vere e quali false (per quelle false trovare un controesempio):

1. se $B = 0$ il sistema ha sempre almeno una soluzione;
2. se $n = m + 1$ il sistema ha sempre almeno una soluzione;
3. se $n > m$ il sistema ha sempre ∞^{m-n} soluzioni;
4. se $n < m$ il sistema non ha mai soluzione.

Esercizio 4.6.33. * L'ingegnere gestionale di una fabbrica metalmeccanica deve gestire il magazzino produzione. Nella fabbrica vengono prodotti bulloni, dadi, rondelle e copiglie. Con una barra di ferro si producono 500 bulloni e 2.000 dadi; con un foglio di lamiera si producono 4.000 rondelle e con una matassa di filo 1.000 copiglie. Nella produzione di ciascuno di essi vi sono mediamente le seguenti percentuali di scarto: 0,5% di bulloni e di dadi, 0,01% di rondelle e copiglie. A fine mese verranno venduti 10.000.000 di bulloni e altrettanti di dadi, 15.000.000 di rondelle e 1.000.000 di copiglie. Quale sistema lineare (in forma matriciale) deve impostare l'ingegnere per sapere quanto nuovo materiale acquistare nel frattempo affinché a inizio del mese seguente il magazzino sia pieno ma senza esuberi?

Esercizio 4.6.34. Considerare in $\mathbb{R}^4$, con coordinate x, y, z, t, il sottospazio vettoriale W definito dal seguente sistema lineare omogeneo:

$$\begin{cases} x - y + t = 0 \\ 2x + y + t = 0 \end{cases}.$$

Descrivere la forma generica degli elementi di W e determinarne la dimensione.

Esercizio 4.6.35. Calcolare il rango delle seguenti matrici:

$$\begin{pmatrix} 1 & 1 & -1 & 2 \\ 1 & 0 & 1 & 0 \\ 5 & 3 & -1 & 6 \end{pmatrix}; \begin{pmatrix} 1 & 1 & -2 \\ 1 & 0 & 1 \\ 3 & -1 & 0 \end{pmatrix}; \begin{pmatrix} 13 & 14 & 15 & 16 \\ 23 & 24 & 25 & 26 \\ 33 & 34 & 35 & 36 \\ 43 & 44 & 45 & 46 \end{pmatrix}.$$

Esercizio 4.6.36. * Considerare il sottospazio vettoriale $W \subset \mathbb{R}^3$ definito dalla seguente equazione omogenea: $ax + by + cz = 0$. Determinare i possibili valori di a, b, c, sapendo che i vettori $(1, 1, 1)$ e $(0, 1, 0)$ appartengono a W.

Esercizio 4.6.37. Calcolare il rango delle seguenti matrici:

$$\begin{pmatrix} 2 & 2 & 2 & 2 \\ 27 & 27 & 27 & 27 \\ 55 & 55 & 55 & 55 \end{pmatrix}; \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}; \begin{pmatrix} 1 & 11 & 111 & 1111 \\ 0 & 10 & 110 & 1110 \\ 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

Esercizio 4.6.38. Si consideri il seguente sistema lineare:

$$\begin{cases} x + y - z + t = 0 \\ x - y - z = 0 \\ 3x - 2y + t = 1 \end{cases}.$$

1. Scrivere in forma matriciale il sistema e ridurlo ad un sistema a scala equivalente usando l'algoritmo di Gauss.
2. Scrivere la soluzione generale del sistema, evidenziando le variabili libere.
3. Indicata con A la matrice incompleta (o dei coefficienti) del sistema, trovare una base dello spazio V dei vettori B per i quali $AX = B$ ammette soluzioni.

Esercizio 4.6.39. Disegnare le soluzioni del seguente sistema lineare

$$\begin{cases} x - y + 3z = 0 \\ x + 2y - z = 0 \end{cases}.$$

Esercizio 4.6.40. Disegnare l'intersezione tra il sottospazio U di $\mathbb{R}^4$ definito da $U := \{(x, y, z, t) \in \mathbb{R}^4 \mid z = 0\}$ e l'insieme delle soluzioni del seguente sistema lineare

$$\begin{cases} x - 2y - z = 2 \\ -x + z + t = 1 \end{cases}.$$

CAPITOLO 5

Applicazioni lineari

Prerequisiti: Conoscenza delle funzioni fra insiemi (Capitolo 1); gli spazi vettoriali (basi, dimensione sottospazi) introdotti nel Capitolo 2; le generalità sulle matrici e le loro operazioni (Capitolo 3); sistemi lineari omogenei e teorema di Rouché–Capelli (Capitolo 4).

Obiettivi: Riconoscere funzioni tra spazi vettoriali che siano applicazioni lineari. Associare una matrice a un'applicazione lineare; saper determinare se un'applicazione lineare fra spazi vettoriali sia suriettiva e/o iniettiva, saper determinare $\ker f$, $\mathrm{Im} f$ per un'applicazione lineare f.

5.1 Introduzione

Abbiamo studiato, nel Capitolo 2, gli spazi vettoriali e ne abbiamo inquadrato vari aspetti (basi, sottospazi, dimensione). Il passo successivo, dopo quanto visto, è quello di considerare le funzioni fra di essi.

Esempio 5.1.1. Consideriamo la seguente funzione $f : \mathbb{R}^2 \to \mathbb{R}^2$, definita da $f(x, y) = (0,8x + 0,1y;\ 0,2x + 0,9y)$. Se riconsideriamo l'esempio 3.3.4, si vede che questa funzione rappresenta i flussi migratori ivi descritti, infatti se $(x, y) =$ (n. abitanti Toscana, n. abitanti resto d'Italia), allora $f(x, y)$ mi dà la distribuzione Toscana-resto d'Italia dopo un anno.

Fra tutte le possibili funzioni, ci interessano in particolare quelle che "rispettano le operazioni" (somma e prodotto per gli scalari) degli spazi vettoriali che sono il loro dominio e codominio. Tali funzioni saranno quelle "più naturali" da utilizzare e sono denominate "applicazioni lineari". Ad esempio, vedremo in seguito come la funzione dei flussi migratori annuali vista nell'esempio precedente sia in effetti un'applicazione lineare.

In questo capitolo si studieranno quindi le funzioni $f : V \to W$ fra due spazi vettoriali V, W. Definiremo che cosa si intende dicendo che la f

è un'applicazione lineare e vedremo che ci sono due sottospazi ($\ker f \subset V$ e $\mathrm{Im} f \subset W$) tali che lo studio delle loro dimensioni ci permette di determinare le proprietà essenziali dell'applicazione lineare f (iniettività e suriettività).

Al fine poi di studiare in pratica le applicazioni lineari, si vedrà come lo strumento ideale per rappresentare queste applicazioni siano le matrici (quando si tratti di spazi vettoriali di dimensione finita, che è il caso su cui ci concentreremo), e come la determinazione delle dimensioni di $\ker f$ e $\mathrm{Im} f$ si possa ridurre a quella dell'individuazione del rango di un'appropriata matrice.

5.2 Applicazioni lineari fra spazi vettoriali

Se si ha una funzione $f : V \to W$ fra due spazi vettoriali V e W, cosa vuol dire che la f "conserva le operazioni" di V e W? Vediamolo per la somma: per ogni $\mathbf{v}_1$ e $\mathbf{v}_2$ in V si dovrà avere che eseguire la loro somma (in V) e poi applicare al risultato la f è la stessa cosa che applicare la f separatamente ai due elementi e poi sommare (in W, ove la f li ha portati); cioè:

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2).$$

Analogamente, per il prodotto per uno scalare, si dovrà avere che se si ha $\mathbf{v} \in V$ e un numero $\alpha \in \mathbb{R}$, moltiplicare $\mathbf{v}$ per α (in V) e poi applicare la f è la stessa cosa che prima applicare la f a $\mathbf{v}$, e poi moltiplicare per α (in W):

$$f(\alpha \mathbf{v}) = \alpha f(\mathbf{v}).$$

Ciò ci conduce alla seguente

Definizione 5.2.1. Siano V, W due spazi vettoriali, e $f : V \to W$ una funzione. Essa si dice un'*applicazione lineare* se verifica le seguenti proprietà:

1. $\forall \mathbf{v}, \mathbf{w} \in V, \qquad f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})$;
2. $\forall \mathbf{v} \in V, \forall \lambda \in \mathbb{R}, \qquad f(\lambda \mathbf{v}) = \lambda f(\mathbf{v})$.

In particolare, si dice un *endomorfismo* un'applicazione lineare da uno spazio vettoriale a se stesso, $f : V \to V$.

Osservazione 5.2.1. Notiamo che come conseguenza immediata della 2), ponendo $\lambda = 0$, si ha:

$$f(\mathbf{0}_V) = \mathbf{0}_W,$$

cioè tutte le applicazioni lineari portano l'elemento nullo dello spazio di partenza nell'elemento nullo di quello d'arrivo.

Esempio 5.2.1. Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$ definita da $f(x, y) = (x+y, x-y)$. È lineare? Affinché lo sia dobbiamo avere $\forall \mathbf{v}_1 = (a, b), \mathbf{v}_2 = (c, d) \in \mathbb{R}^2$ e $\lambda \in \mathbb{R}$ devono valere:

1. $f((a, b) + (c, d)) = f(a, b) + f(c, d)$,
2. $f(\lambda(a, b)) = \lambda f(a, b)$.

Andiamo a controllare le due uguaglianze:

1. $f((a, b) + (c, d)) = f(a + c, b + d) = (a + c + b + d, a + c - b - d)$;
   $f(a, b) + f(c, d) = (a + b, a - b) + (c + d, c - d) =$
   $= (a + b + c + d, a - b + c - d)$;
2. $f(\lambda(a, b)) = f(\lambda a, \lambda b) = (\lambda a + \lambda b, \lambda a - \lambda b) = \lambda(a + b, a - b) ==$
   $\lambda f(a, b)$.

Quindi la f è un'applicazione lineare.

Esempio 5.2.2. Qualunque siano gli spazi vettoriali V,W, la *funzione nulla*: $f(\mathbf{v}) = \mathbf{0}_W$, $\forall\, \mathbf{v} \in V$; è banalmente lineare.

Per ogni spazio vettoriale V la funzione identità $id_V : V \to V$, ove $id_V(\mathbf{v}) = \mathbf{v}$, $\forall\, \mathbf{v} \in V$ è anch'essa ovviamente lineare.

Esempio 5.2.3. Sia $f : \mathbb{R}^3 \to \mathbb{R}^2$ definita da $f(x, y, z) = (x + y, z)$. Per vedere se f sia lineare o meno consideriamo $(x, y, z), (t, u, v) \in \mathbb{R}^3$ e $\lambda \in \mathbb{R}$; avremo

$$f((x, y, z) + (t, u, v)) = f(x + t, y + u, z + v) = (x + t + y + u, z + v) =$$

$$= (x + y, z) + (t + u, v) = f(x, y, z) + f(t, u, v);$$

$$f(\lambda(x, y, z)) = f(\lambda x, \lambda y, \lambda z) = (\lambda x + \lambda y, \lambda z) = \lambda(x + y, z) = \lambda f(x, y, z),$$

quindi la f è un'applicazione lineare.

Esempio 5.2.4. Sia ora $f : \mathbb{R}^2 \to \mathbb{R}^2$, con $f(x, y) = (x - y, 1)$.
Abbiamo $f(0, 0) = (0, 1)$, quindi f non è lineare (non manda il vettore nullo di $\mathbb{R}^2$ in se stesso, come dovrebbe).

Esempio 5.2.5. Data $f : \mathbb{R}^2 \to \mathbb{R}^2$, con $f(x, y) = (x^2, y)$, si ha: $f(\lambda x, \lambda y) = (\lambda^2 x^2, \lambda y)$, mentre $\lambda f(x, y) = (\lambda x^2, \lambda y)$; quindi la f non è lineare.

Esempio 5.2.6. Data $f : \mathbb{R} \to \mathbb{R}$, se vogliamo che sia lineare, dovremo avere che $f(\alpha x) = \alpha f(x)$ e cioè che la quantità $f(x)$ sia direttamente proporzionale a x; questo ci dice che nel caso delle applicazioni da $\mathbb{R}$ in sé, quelle lineari possono solo essere quelle per cui fra x e $y = f(x)$ c'è una proporzionalità diretta. È ben noto che tali funzioni sono del tipo $f(x) = kx$, per ogni $k \in \mathbb{R}$. All'inverso è facile vedere che una tale funzione è lineare (farlo per esercizio), quindi si avrà che:

Le sole applicazioni lineari $f : \mathbb{R} \to \mathbb{R}$ sono le proporzionalità dirette; cioè tutte e sole le funzioni $f(x) = kx$, $k \in \mathbb{R}$.

Esempio 5.2.7. Sia ora $f : \mathbb{R}[x] \to \mathbb{R}^2$, con
$f(a_0 + a_1x + \cdots + a_sx^s) = (a_0, a_1)$.
Per vedere se f sia lineare, consideriamo $p(x) = a_0 + a_1x + \cdots + a_sx^s$ e $q(x) = b_0 + b_1x + \cdots + b_tx^t$ e supponiamo $s \leq t$. Avremo:

$$f(p(x)+q(x)) = f((a_0+b_0)+(a_1+b_1)x+\cdots+(a_s+b_s)x^s+\cdots+b_tx^t) =$$

$$= ((a_0+b_0),(a_1+b_1)) = (a_0,a_1)+(b_0,b_1) = f(p(x))+f(q(x)).$$

Verificare per esercizio che anche $f(\lambda p(x)) = \lambda f(p(x))$, e che quindi f è lineare.

Osservazione 5.2.2. Consideriamo una matrice $A \in \mathbb{R}^{m,n}$; vediamo che è possibile associare a essa un'applicazione lineare $f_A : \mathbb{R}^n \to \mathbb{R}^m$. Procederemo in questo modo: ogni elemento $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ può essere visto come una matrice colonna in $\mathbb{R}^{n,1}$:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

Si definisce allora $f_A(X) = A \cdot X \in \mathbb{R}^{m,1}$ (che identifichiamo con $\mathbb{R}^m$). Si verifica immediatamente che:
$\forall X_1, X_2, X \in \mathbb{R}^n$, $\forall \lambda \in \mathbb{R}$:

$$f_A(X_1 + X_2) = A(X_1 + X_2) = AX_1 + AX_2 = f_A(X_1) + f_A(X_2);$$

$$f_A(\lambda X) = A(\lambda X) = \lambda AX = \lambda f_A(X),$$

quindi f_A è un'applicazione lineare.

Questo ci dice che la funzione $f : \mathbb{R}^2 \to \mathbb{R}^2$ vista nell'Esempio 5.2.1 può esser definita usando la matrice dei flussi migratori dell'Esempio 3.3.5, e quindi è un'applicazione lineare. Vediamo un altro esempio;

Esempio 5.2.8. Sia $A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{2,3}$; allora

$$f_A(x,y,z) = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+2y+z \\ x+y \end{pmatrix} \in \mathbb{R}^{2,1}$$

quindi la funzione associata alla A (identificando $\mathbb{R}^{2,1}$ con $\mathbb{R}^2$) è :
$f_A : \mathbb{R}^3 \to \mathbb{R}^2$, con $f_A(x,y,z) \to (2x+2y+z, x+y)$.

5.3 Nucleo e immagine, iniettività e suriettività

Vediamo adesso due sottospazi vettoriali che sono "naturalmente" associati a ogni applicazione lineare, e tali che le loro dimensioni ci danno informazioni sostanziali sulla f.

Definizione 5.3.1. Data una applicazione lineare $f : V \to W$ consideriamo i due insiemi associati a essa:
L'*immagine di* f: $\quad \mathrm{Im} f = \{\mathbf{w} \in W \mid \exists\, \mathbf{v} \in V, f(\mathbf{v}) = \mathbf{w}\}$;
Il *nucleo di* f: $\quad \ker f = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0}_W\}$.

Quindi i vettori di $\mathrm{Im}\ f$ sono quei vettori $\mathbf{w}$ di W che provengono, via f, da qualche vettore $\mathbf{v}$ di $\mathbb{R}^n$, mentre i vettori di $\ker f$ sono quei vettori di V che la f porta nel vettore nullo di W.

Non è difficile vedere che $\mathrm{Im}\ f$ e $\ker f$ sono sottospazi, rispettivamente, di W e di V.

Proposizione 5.3.1. [$\mathrm{Im} f$ e $\ker f$ sono sottospazi] *Sia $f : V \to W$ un'applicazione lineare fra gli spazi vettoriali V e W. Allora $\ker f$ è un sottospazio di V e $\mathrm{Im} f$ è un sottospazio di W.*

Dimostrazione. Consideriamo per primo $\ker f$; banalmente si ha che $\mathbf{0}_V \in \ker f$ e $\forall\, \mathbf{v}, \mathbf{v}' \in \ker f$, $\forall\, \alpha, \beta \in \mathbb{R}$:

$$f(\alpha\mathbf{v} + \beta\mathbf{v}') = f(\alpha\mathbf{v}) + f(\beta\mathbf{v}') = \alpha f(\mathbf{v}) + \beta f(\mathbf{v}') = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W.$$

quindi $\alpha\mathbf{v} + \beta\mathbf{v}' \in \ker f$, e $\ker f$ risulta un sottospazio di V.

Per $\mathrm{Im} f$: se $\mathbf{w}, \mathbf{w}' \in \mathrm{Im} f$, allora: $\exists\, \mathbf{v}, \mathbf{v}' \in V$; tali che $f(\mathbf{v}) = \mathbf{w}$; $f(\mathbf{v}') = \mathbf{w}'$. Allora, $\forall\, \alpha, \beta \in \mathbb{R}$:

$$f(\alpha\mathbf{v} + \beta\mathbf{v}') = f(\alpha\mathbf{v}) + f(\beta\mathbf{v}') = \alpha f(\mathbf{v}) + \beta f(\mathbf{v}') = \alpha\mathbf{w} + \beta\mathbf{w}',$$

quindi $\alpha\mathbf{w} + \beta\mathbf{w}' \in \mathrm{Im} f$, e poiché banalmente si ha $\mathbf{0}_W \in \mathrm{Im} f$ in quanto $f(\mathbf{0}_V) = \mathbf{0}_W$, $\mathrm{Im} f$ risulta un sottospazio di W. □

Esempio 5.3.1. Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$, con $f(x, y, z) = (x+y, y+z, x+2y+z)$. La f è lineare (verificarlo per esercizio) e si ha:

$$\ker f = \{(x, y, z) \in \mathbb{R}^3 | f(x, y, z) = (0, 0, 0)\} = \\ = \{(x, y, z) \in \mathbb{R}^3 | x + y = 0;\ y + z = 0;\ x + 2y + z = 0\};$$

$$\text{Im } f = \{(a,b,c) \in \mathbb{R}^3 | \exists\, (x,y,z) \in \mathbb{R}^3,\, f(x,y,z) = (a,b,c)\} =$$
$$= \{(a,b,c) \in \mathbb{R}^3 | \exists\, (x,y,z) \in \mathbb{R}^3, x+y=a;\ y+z=b; x+2y+z=c\}.$$

Vediamo di determinare più precisamente $\text{Im} f$ e $\ker f$, e darne una base come sottospazi vettoriali.

Per determinare $\ker f$ si deve risolvere il sistema:

$$\begin{cases} x+y=0 \\ y+z=0 \\ x+2y+z=0 \end{cases} \rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Cioè $\ker f$ non è altro che lo spazio delle soluzioni di tale sistema (omogeneo). L'eliminazione di Gauss sulla matrice porta:

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

quindi la matrice dei coefficienti ha rango 2, e il sistema ha $\infty^{3-2} = \infty^1$ soluzioni, perciò $\dim(\ker f) = 1$. Se vogliamo una base di $\ker f$, si risolve il sistema ottenendo: $y = -z$, $x = z$, quindi $\ker f = \{(z,-z,z)\}$, e una base è $\{(1,-1,1)\}$.

Per determinare $\text{Im} f$ si deve risolvere il sistema:

$$\begin{cases} x+y=a \\ y+z=b \\ x+2y+z=c \end{cases} \rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

perché il sistema abbia soluzioni si deve avere (dal teorema di Rouché-Capelli, vedi Capitolo 4)

$$r \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} = r \begin{pmatrix} 1 & 1 & 0 & a \\ 0 & 1 & 1 & b \\ 1 & 2 & 1 & c \end{pmatrix}$$

ossia, affinché $(a,b,c) \in \text{Im} f$, si deve avere che (a,b,c) appartiene allo spazio dalle colonne di A, $\mathcal{C}(A)$, ove A è la matrice dei coefficienti del sistema, vale a dire che (a,b,c) deve essere generato dalle colonne di A. Abbiamo visto prima che $r(A) = 2$, quindi per avere una base di $\mathcal{C}(A)$ basta prendere due colonne indipendenti di A; scegliendo le prime due avremo:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \alpha+\beta \\ \beta \\ \alpha+2\beta \end{pmatrix},$$

quindi gli elementi di $\text{Im} f$ sono tutti i vettori del tipo $(\alpha+\beta, \beta, \alpha+2\beta)$, al variare di $\alpha, \beta \in \mathrm{R}$:

$$\text{Im} f = \{(a,b,c) \in \mathbb{R}^3 \,|\, \exists\, \alpha, \beta \in \mathrm{R}, (a,b,c) = (\alpha+\beta, \beta, \alpha+2\beta)\}.$$

La seguente proposizione ci mostrerà come si possono in generale trovare dei generatori per $\mathrm{Im}f$.

Proposizione 5.3.2. *Sia $f : V \to W$, lineare, e sia $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base di V. Allora $\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)\}$ è un insieme di generatori per $\mathrm{Im}f$.*

Dimostrazione. La dimostrazione è piuttosto immediata. Se abbiamo un qualsiasi $\mathbf{w} \in \mathrm{Im}f$, allora esiste $\mathbf{v} \in V$ tale che $f(\mathbf{v}) = \mathbf{w}$. Scrivendo $\mathbf{v}$ tramite la base $\mathcal{B}$, si avrà: $\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$, e quindi anche $\mathbf{w} = f(\mathbf{v}) = f(\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n) = \alpha_1 f(\mathbf{v}_1) + \cdots + \alpha_n f(\mathbf{v}_n)$. Perciò i vettori $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)$ generano ogni vettore di $\mathrm{Im}f$, costituendo un suo sistema di generatori. □

Osservazione 5.3.1. Non è detto che $\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)\}$ sia una base di $\mathrm{Im}f$, in quanto i vettori $f(\mathbf{v}_i)$ potrebbero non essere indipendenti (anche se i $\mathbf{v}_i$ lo sono); una base si otterrà scegliendo il massimo numero possibile di vettori indipendenti nell'insieme di generatori $\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)\}$; tale numero ci darà quindi il valore di $\dim(\mathrm{Im}f)$.

Esempio 5.3.2. Riprendiamo l'esempio precedente, ove $f : \mathbb{R}^3 \to \mathbb{R}^3$, con $f(x, y, z) = (x+y, y+z, x+2y+z)$. Considerando la base canonica in $\mathbb{R}^3$, si ha: $f(\mathbf{1}, 0, 0) = (1, 0, \mathbf{1})$, $f(0, 1, 0) = (1, \mathbf{1}, 2)$, $f(0, 0, 1) = (0, 1, \mathbf{1})$. Quindi i vettori $(1, 0, \mathbf{1})$, $(\mathbf{1}, 1, 2)$, $(0, \mathbf{1}, \mathbf{1})$ sono un insieme di generatori per $\mathrm{Im}f$. Notiamo che non sono una base, in quanto abbiamo visto che qui $\dim(\mathrm{Im}f) = \mathbf{2}$. Se scegliamo i primi due (come si è fatto precedentemente), avremo che $\{(\mathbf{1}, 0, 1), (1, \mathbf{1}, 2)\}$ è una base di $\mathrm{Im}f$, infatti ogni $\mathbf{w} \in \mathrm{Im}f$ si può scrivere:

$$(\alpha + \beta, \beta, \alpha + 2\beta) = \alpha(1, 0, 1) + \beta(1, 1, 2).$$

Vediamo adesso come la conoscenza di $\mathrm{Im}f$ e $\ker f$ ci dia informazioni su iniettività e suriettività della f (concetti introdotti nel Capitolo 1).

Proposizione 5.3.3. *Sia $f : V \to W$ un'applicazione lineare; allora*

- *f è suriettiva se e solo se $\mathrm{Im}f = W$;*
- *f è iniettiva se e solo se $\ker f = \{\mathbf{0}_V\}$.*

Dimostrazione. Il primo punto è piuttosto immediato, in quanto dire che la f è suriettiva significa proprio che ogni vettore di W proviene, via f, da uno di V, cioè che $f(V) = W$, vale a dire $\mathrm{Im}f = W$.

Vediamo adesso il secondo punto; dovremo verificare due implicazioni:

"f iniettiva $\Rightarrow \ker f = \{\mathbf{0}_V\}$" e "$\ker f = \{\mathbf{0}_V\} \Rightarrow f$ iniettiva".

Partiamo dalla più semplice: Supponiamo che f sia iniettiva; allora, poiché $f(\mathbf{0}_V) = \mathbf{0}_W$, nessun altro $\mathbf{v} \in V$ è tale che $f(\mathbf{v}) = \mathbf{0}_W$, e quindi solo lo $\mathbf{0}_V$ andrà nello $\mathbf{0}_W$, cioè $\ker f = \{\mathbf{0}_V\}$.

Supponiamo invece $\ker f = \{\mathbf{0}_V\}$, e consideriamo $\mathbf{v}, \mathbf{v}' \in V$ e tali che $f(\mathbf{v}) = f(\mathbf{v}')$. Allora avremo $f(\mathbf{v}) - f(\mathbf{v}') = \mathbf{0}_W$, indi $f(\mathbf{v} - \mathbf{v}') = \mathbf{0}_W$, e allora $\mathbf{v} - \mathbf{v}' \in \ker f$.

Ma $\ker f = \{\mathbf{0}_V\}$, quindi $\mathbf{v} - \mathbf{v}' = \mathbf{0}_V$, cioè $\mathbf{v} = \mathbf{v}'$ e la f risulta iniettiva (solo vettori uguali hanno immagine uguale). □

Si ha quindi che f è suriettiva se la sua immagine è "massima", e che è iniettiva se il suo nucleo è "minimo".

Le dimensioni di $\mathrm{Im} f$ e $\ker f$ non sono indipendenti fra di loro; abbiamo infatti la seguente

Proposizione 5.3.4. *Sia $f : V \to W$ un'applicazione lineare fra gli spazi vettoriali V e W, con $\dim V = n$. Allora vale:*

$$\dim(\mathrm{Im} f) + \dim(\ker f) = n.$$

Dimostrazione. Sia $\dim(\ker f) = r \leq n$, e sia $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ una base di $\ker f$. Sappiamo che si può completarla a una base di V:

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}.$$

Se dimostriamo che gli $n - r$ vettori che abbiamo aggiunto sono tali che $\{f(\mathbf{v}_{r+1}), \ldots, f(\mathbf{v}_n)\}$ è una base di $\mathrm{Im} f$, abbiamo terminato, in quanto avremo $\dim(\mathrm{Im} f) + \dim(\ker f) = r + (n - r) = n = \dim V$.

Dalla Proposizione 5.3.2 sappiamo che l'insieme

$$\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_r), f(\mathbf{v}_{r+1}), \ldots, f(\mathbf{v}_n)\}$$

è un insieme di generatori per $\mathrm{Im} f$; ma $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \in \ker f$ e quindi $f(\mathbf{v}_1) = \ldots = f(\mathbf{v}_r) = \mathbf{0}_W$, allora già $\{f(\mathbf{v}_{r+1}), \ldots, f(\mathbf{v}_n)\}$ è un insieme di generatori per $\mathrm{Im} f$. Ci resta da vedere che siano linearmente indipendenti. Consideriamo una combinazione lineare:

$a_{r+1} f(\mathbf{v}_{r+1}) + \cdots + a_n f(\mathbf{v}_n) = \mathbf{0}_W$,

dobbiamo avere che gli a_i sono tutti nulli.

Si ha (per la linearità della f): $f(a_{r+1}\mathbf{v}_{r+1} + \cdots + a_n\mathbf{v}_n) = \mathbf{0}_W$; e quindi $a_{r+1}\mathbf{v}_{r+1} + \cdots + a_n\mathbf{v}_n \in \ker f$. Ma poiché $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ è una base di $\ker f$, il vettore $a_{r+1}\mathbf{v}_{r+1} + \cdots + a_n\mathbf{v}_n$ si può esprimere come combinazione lineare di $\mathbf{v}_1, \ldots, \mathbf{v}_r$:

$a_{r+1}\mathbf{v}_{r+1} + \cdots + a_n\mathbf{v}_n = a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r$. E così abbiamo:

$$a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r - a_{r+1}\mathbf{v}_{r+1} - \cdots - a_n\mathbf{v}_n = \mathbf{0}_V,$$

ma $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ sono linearmente indipendenti (formano una base) e quindi gli a_i, $i = 1, \ldots, n$ sono tutti nulli. Da ciò segue che

$\{f(\mathbf{v}_{r+1}), \ldots, f(\mathbf{v}_n)\}$ sono linearmente indipendenti, che è quello che volevamo. □

Osservazione 5.3.2. La precedente proposizione vale anche nel caso infinito, cioè $\dim(\mathrm{Im} f) + \dim(\ker f) = \dim V$ anche quando $\dim V = \infty$ (in quel caso si dovrà considerare $\infty + r = \infty$ o $\infty + \infty = \infty$).

Dimostrazione. Se $\dim V = \infty$ e anche $\dim(\ker f) = \infty$, l'enunciato vale ovviamente (qualsiasi valore abbia $\dim(\mathrm{Im} f)$). Consideriamo il caso in cui $\dim V = \infty$ e $\dim(\ker f) = r$. Procedendo come nel caso finito, si avrà che il completamento della base $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ di $\ker f$ a una base di V è dato da un insieme di infiniti vettori, $\{\mathbf{v}_\alpha\}_{\alpha \in K}$, le cui immagini generano $\mathrm{Im} f$. Per vedere che i vettori $f(\mathbf{v}_\alpha)$ sono linearmente indipendenti, basta vedere che se, per assurdo, ci fossero fra di essi dei vettori linearmente dipendenti, si avrebbe una combinazione lineare $a_{\alpha_1} f(\mathbf{v}_{\alpha_1}) + \cdots + a_{\alpha_t} f(\mathbf{v}_{\alpha_t}) = \mathbf{0}_W$, con gli a_{α_j} (ove $\alpha_j \in K$) non tutti nulli. Come nel caso finito, ciò ci darebbe che $a_{\alpha_1}\mathbf{v}_{\alpha_1} + \cdots + a_{\alpha_t}\mathbf{v}_{\alpha_t} \in \ker f$, e quindi che c'è una combinazione lineare $a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r - a_{\alpha_1}\mathbf{v}_{\alpha_1} - \cdots - a_{\alpha_t}\mathbf{v}_{\alpha_t} = \mathbf{0}_V$ a coefficienti non tutti nulli, il che è assurdo perché questi vettori fanno parte di una base (infinita). □

Una conseguenza della Proposizione 5.3.4 è la seguente:

Proposizione 5.3.5. *Sia $f : V \to W$, con $\dim V = n = \dim W$. Allora la f è iniettiva se e solo se è suriettiva.*

Dimostrazione. Poichè $\dim(\mathrm{Im} f) + \dim(\ker f) = n$, si ha che f è suriettiva, cioè $\dim(\mathrm{Im} f) = n$, se e solo se $\dim(\ker f) = 0$; cioè se e solo se f è iniettiva. □

Notiamo che la condizione $\dim V = \dim W$ nella proposizione precedente è essenziale; infatti se essa non vale allora iniettività e suriettività sono indipendenti; ad esempio la $f : \mathbb{R}^3 \to \mathbb{R}^2$, con $f(x, y, z) = (x, y)$, è suriettiva ma non iniettiva (infatti $\ker f = \{(0, 0, z) | z \in \mathbb{R}\}$).

Notiamo anche che il risultato non vale se $\dim V = \dim W = \infty$, come mostrato dal seguente

Esempio 5.3.3. Sia $D : \mathbb{R}[x] \to \mathbb{R}[x]$, ove $D(p(x)) = p'(x)$ (la D è la funzione "derivata"). Si ha che la D è lineare (vedi Esercizio 5.6.22), non è iniettiva (se due polinomi $p(x)$ e $q(x)$ differiscono solo per il termine noto allora $D(p(x)) = D(q(x))$ perché hanno la stessa derivata). Però la D è suriettiva (ogni polinomio è integrabile, quindi esistono sue primitive).

Osserviamo che da quanto visto finora si può ricavare il seguente risultato:

Proposizione 5.3.6. *Sia $f : V \to W$ un'applicazione lineare e siano* $\dim V = n$, $\dim W = m$. *Allora:*

1. *Se $n < m$ la f non può essere suriettiva.*
2. *Se $n > m$ la f non può essere iniettiva.*

Dimostrazione. 1) È noto che se i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ formano una base per V, allora $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)$ sono generatori per $\mathrm{Im}f$, quindi $\dim(\mathrm{Im}f) \leq n$, e se $n < m$, non si potrà avere $\mathrm{Im}f = W$, cioè la f non potrà essere suriettiva.

2) f è iniettiva se e solo se $\ker f = \{\mathbf{0}_V\}$, cioè $\dim(\ker f) = 0$. Ciò risulta impossibile poiché

$$\dim(\ker f) = n - \dim(\mathrm{Im}f) \geq n - m > 0.$$

□

Esempio 5.3.4. Sia $f : \mathbb{R}^2 \to \mathbb{R}^5$, con $f(x, y) = (x, y, x, y, x + y)$. È facile vedere che la f è lineare (farlo per esercizio); la f è suriettiva e/o iniettiva? Dalla proposizione precedente si ha che la f non può essere suriettiva (possiamo anche notare che in $\mathrm{Im}f$ ci son solo vettori con la prima coordinata uguale alla terza e la seconda alla quarta, quindi $\mathrm{Im}f \neq \mathbb{R}^5$ in quanto contiene solo quel tipo di vettori).

Per l'iniettività abbiamo che $(x, y) \in \ker f$ se e solo se $f(x, y) = \mathbf{0}$, e cioè se $x = 0, y = 0, x = \mathbf{0}, y = 0, x + y = 0$, quindi $\ker f = \{(0, 0)\}$ e f è iniettiva.

Un caso che ci interessa particolarmente è quello in cui la f sia contemporaneamente iniettiva e suriettiva, cioè che sia biunivoca.

Definizione 5.3.2. Una $f : V \to W$ si dice un *isomorfismo* se è biunivoca; in questo caso V e W si dicono *isomorfi* e si scrive $V \cong W$.

Gli isomorfismi sono particolarmente importanti quando si studiano gli spazi vettoriali; ad esempio quando abbiamo considerato gli spazi $\mathcal{V}_O^2$, $\mathcal{V}_O^3$ e abbiamo notato come il loro studio si possa eseguire tramite un "passaggio" a $\mathbb{R}^2$ e $\mathbb{R}^3$ utilizzando un sistema di coordinate cartesiane; quello che in effetti abbiamo fatto è di definire un isomorfismo tra i due spazi, $f : \mathcal{V}_O^i \to \mathbb{R}^i$, $i = 2, 3$, che costituisce una specie di "dizionario" fra i due spazi da comparare.

Questo procedimento è generale: se abbiamo un qualsiasi spazio vettoriale V, con $\dim V = n > 0$, fissando una base $\mathcal{B}$ di V avremo che a ogni suo vettore $\mathbf{v} \in V$ rimane assegnata una n-pla di numeri reali (le sue coordinate rispetto alla base $\mathcal{B}$), e quindi un elemento di $\mathbb{R}^n$;

tale associazione ci permette di studiare V "come se fosse $\mathbb{R}^n$". Più precisamente:

Proposizione 5.3.7. *Sia V uno spazio vettoriale non nullo e tale che $\dim V = n$. Allora V e $\mathbb{R}^n$ sono isomorfi.*

Dimostrazione. Dobbiamo costruire un isomorfismo $V \to \mathbb{R}^n$; fissiamo una base $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ di V e consideriamo la funzione $x_{\mathcal{B}} : V \to \mathbb{R}^n$ definita nel Capitolo 2, ove per ogni $\mathbf{v} \in V$, $x_{\mathcal{B}}(\mathbf{v})$ è l'unica n-pla di numeri reali $(x_1, \ldots, x_n)$ costituita dalle sue coordinate rispetto alla base $\mathcal{B}$, cioè tali che: $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$. La $x_{\mathcal{B}}$ sarà iniettiva perché $\ker x_{\mathcal{B}}$ è dato dai vettori che hanno coordinate $(0, \ldots, 0)$, e cioè dal solo vettore nullo. Sarà anche suriettiva, in quanto ogni n-pla $(x_1, \ldots, x_n) \in \mathbb{R}^n$ proviene, tramie la $x_{\mathcal{B}}$, da un vettore di V (il vettore $x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$).

Quindi la f è un isomorfismo. □

Osservazione 5.3.3. Sia $f : V \to W$ un isomorfismo; essendo f biunivoca potremo definire l'applicazione $f^{-1} : W \to V$, con $\forall \mathbf{w} \in W$: $f^{-1}(\mathbf{w}) = \mathbf{v}$, ove $\mathbf{v}$ è l'unico vettore di V tale che $f(\mathbf{v}) = \mathbf{w}$ (vedi anche Capitolo 1).

La f^{-1} è tale che

$$f^{-1} \circ f = id_V; \qquad f \circ f^{-1} = id_W$$

ove id_V e id_W sono le applicazioni identità in V e W.

Proposizione 5.3.8. *Sia $f : V \to W$, un isomorfismo. Si consideri la funzione $f^{-1} : W \to V$: anch'essa risulta lineare (e quindi un isomorfismo).*

Dimostrazione. Dobbiamo vedere che $\forall \mathbf{w}_1, \mathbf{w}_2 \in W$,

$f^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = f^{-1}(\mathbf{w}_1) + f^{-1}(\mathbf{w}_2)$

e che $\forall \mathbf{w} \in W, \forall \lambda \in \mathbb{R}$, $f^{-1}(\lambda\mathbf{w}) = \lambda f^{-1}(\mathbf{w})$.

Per la prima cosa, sia $f^{-1}(\mathbf{w}_1) = \mathbf{v}_1$, $f^{-1}(\mathbf{w}_2) = \mathbf{v}_2$; allora si avrà :

$$f^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = f^{-1}(f(\mathbf{v}_1) + f(\mathbf{v}_2)) = f^{-1}(f(\mathbf{v}_1 + \mathbf{v}_2)) =$$

$$= \mathbf{v}_1 + \mathbf{v}_2 = f^{-1}(\mathbf{w}_1) + f^{-1}(\mathbf{w}_2).$$

Per la seconda proprietà , avremo:

$$f(\lambda f^{-1}(\mathbf{w})) = \lambda f(f^{-1}(\mathbf{w})) = \lambda\mathbf{w} = f(f^{-1}(\lambda\mathbf{w})).$$

Poiché la f è iniettiva, si ha $\lambda f^{-1}(\mathbf{w}) = f^{-1}(\lambda\mathbf{w})$, come dovevasi dimostrare. □

5.4 Applicazioni lineari e matrici

Abbiamo visto nell'Osservazione 5.2.2 che data una matrice $A \in \mathbb{R}^{m,n}$ è possibile associarle un'applicazione lineare $f_A : \mathbb{R}^n \to \mathbb{R}^m$ e che nell'Esempio 5.3.1 per determinare $\mathrm{Im} f$ e $\ker f$ si lavorava con una matrice A (dovendo risolvere dei sistemi lineari). Ci chiediamo adesso se si possa fare ciò in generale: dati due spazi vettoriali V, W di dimensioni n, m rispettivamente, e un'applicazione lineare $f : V \to W$, vogliamo associare a essa una matrice che contenga tutta " l'informazione " associata alla f.

Si scelgano due basi $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ e $\mathcal{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$, rispettivamente di V e di W; e si considerino poi i vettori: $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n) \in W$ esprimendoli come combinazioni lineari dei $\mathbf{w}_i$:

$$\begin{aligned} f(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m; \\ f(\mathbf{v}_2) &= a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m; \\ &\vdots \\ f(\mathbf{v}_n) &= a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m. \end{aligned}$$

I coefficienti a_{ij} vengono allora a formare una matrice $A = (a_{ij}) \in \mathbb{R}^{m,n}$, che sarà la matrice cercata. Sia infatti

$$\mathbf{v} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n \in V$$

un qualsiasi vettore di V; esso corrisponderà , tramite le sue coordinate rispetto alla base $\mathcal{B}$, al vettore $x_{\mathbf{v}}(\mathcal{B}) = (b_1, \ldots, b_n) \in \mathbb{R}^n$ (vedi Capitolo 2); per l'uso in calcolo matriciale scriveremo $x_{\mathbf{v}}(\mathcal{B})$ come vettore colonna $X \in \mathbb{R}^{n,1}$:

$$X = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}_{\mathcal{B}} \qquad \leftrightarrow \qquad \mathbf{v} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n.$$

Consideriamo adesso il prodotto $A \cdot X$; avremo:

$$A \cdot X = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_n \end{pmatrix} = C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} \in \mathbb{R}^m,$$

ove, $\forall\, i = 1, \ldots, m$, si ha:

$$c_i = \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_{i1}b_1 + \cdots + a_{in}b_n = \sum_{k=1}^{n} a_{ik}b_k.$$

Ma il vettore colonna C può rappresentare, tramite la base $\mathcal{B}'$, un vettore di W:

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}_{\mathcal{B}'} \leftrightarrow c_1\mathbf{w}_1 + \cdots + c_m\mathbf{w}_m;$$

si ha quindi:

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}_{\mathcal{B}'} = (\sum_{k=1}^{n} a_{1k}b_k)\mathbf{w}_1 + (\sum_{k=1}^{n} a_{2k}b_k)\mathbf{w}_2 + \cdots + (\sum_{k=1}^{n} a_{mk}b_k)\mathbf{w}_m.$$

Riordinando tale somma in modo diverso (mettendo in evidenza i coefficienti b_k), si ha:

$$b_1(a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m) + b_2(a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m) + \cdots$$

$$\cdots + b_n(a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m) =$$

$$= b_1 f(\mathbf{v}_1) + \cdots + b_n f(\mathbf{v}_n) = f(b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n) = f(\mathbf{v}).$$

In conclusione si è ottenuto che la matrice A rappresenta la f in quanto il prodotto $A \cdot X$, ci dà le coordinate di $f(\mathbf{v})$ rispetto a $\mathcal{B}'$, una volta che si rappresentino i vettori in V tramite la base $\mathcal{B}$ e quelli di W tramite la base $\mathcal{B}'$.

Esprimeremo questo fatto indicando la matrice A associata alla f con il simbolo $A = M_{\mathcal{B}\mathcal{B}'}(f)$, che segnala come la matrice dipenda dalla scelta delle basi; diremo che $M_{\mathcal{B}\mathcal{B}'}(f)$ è la matrice che rappresenta la f rispetto alle basi $\mathcal{B}$ (in partenza) e $\mathcal{B}'$ (in arrivo).

Si ha così che la matrice associata a una f non è unica, ma dipende dalla scelta delle due basi, sia nello spazio di arrivo che in quello di partenza.

Le colonne della matrice $M_{\mathcal{B}\mathcal{B}'}(f)$ restano determinate dai vettori $f(\mathbf{v}_i)$ (immagini della base di partenza), scritti come vettori colonna espressi rispetto alla base di arrivo:

$$f(\mathbf{v}_1) = a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}_{\mathcal{B}'};$$

$$f(\mathbf{v}_2) = a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}_{\mathcal{B}'};$$

$$\vdots$$

$$f(\mathbf{v}_n) = a_{1n}\mathbf{w_1} + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}_{\mathcal{B}'}.$$

La seguente proposizione riassume quanto visto fin qui:

Proposizione 5.4.1. *Sia $f : V \to W$ un'applicazione lineare, e siano V,W spazi vettoriali di dimensioni (finite) $\dim V = n$, $\dim W = m$. Date comunque due basi $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $\mathcal{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$, rispettivamente di V e di W, se $\forall\, j = 1, \ldots, n$ si ha $f(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + \cdots + a_{mj}\mathbf{w}_m$, allora la matrice $M_{\mathcal{B},\mathcal{B}'}(f) = (a_{i,j}) \in \mathbb{R}^{m,n}$ è tale che $\forall\, \mathbf{v} \in V$, il vettore $f(\mathbf{v}) \in W$ ha coordinate rispetto a $\mathcal{B}'$:*

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}_{\mathcal{B}'} = M_{\mathcal{B},\mathcal{B}'}(f) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}_{\mathcal{B}}$$

dove $(b_1, \ldots, b_n)$ sono le coordinate di v rispetto a $\mathcal{B}$.

Esempio 5.4.1. Riprendiamo l'applicazione lineare considerata precedentemente: $f : \mathbb{R}^3 \to \mathbb{R}^2$, definita da $f(x, y, z) = (x + y, z)$. Consideriamo le basi canoniche $\mathcal{E}_3$ e $\mathcal{E}_2$ nei due spazi, e cerchiamo di determinare $M_{\mathcal{E}_3\mathcal{E}_2}(f)$.

Abbiamo: $f(1,0,0) = (\mathbf{1},\mathbf{0})$; $f(\mathbf{0},\mathbf{1},\mathbf{0}) = (1,0)$ e $f(\mathbf{0},\mathbf{0},\mathbf{1}) = (0,1)$. Allora la matrice cercata risulta essere:

$$M_{\mathcal{E}_3\mathcal{E}_2}(f) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Infatti, $\forall\, (x, y, z) \in \mathbb{R}^3$, avremo:

$$f(x,y,z) = xf(\mathbf{1},0,0) + yf(\mathbf{0},\mathbf{1},\mathbf{0}) + zf(0,0,\mathbf{1}) =$$

$$= x\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{E}_2} + y\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{E}_2} + z\begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{E}_2} =$$

$$= \begin{pmatrix} x+y \\ z \end{pmatrix}_{\mathcal{E}_2} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{E}_3}.$$

Osservazione 5.4.1. Come si vede dall'esempio, quando si tratta con gli spazi $\mathbb{R}^n$, $\mathbb{R}^m$, l'uso delle basi canoniche $\mathcal{E}_n$, $\mathcal{E}_m$ rende molto più semplice la determinazione di una matrice da associare a $f : \mathbb{R}^n \to \mathbb{R}^m$, in quanto le immagini $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$ dei vettori della base di partenza danno già le colonne della matrice cercata.

Esempio 5.4.2. Riprendiamo ancora l'applicazione lineare appena considerata:

$f : \mathbb{R}^3 \to \mathbb{R}^2$, definita da $f(x, y, z) = (x + y, z)$. Consideriamo stavolta, invece delle basi canoniche, le basi $\mathcal{B} = \{(1, 1, 0), (0, 1, 0), (0, 1, 1)\}$ in $\mathbb{R}^3$, e $\mathcal{B}' = \{(1, 1), (1, -1)\}$ in $\mathbb{R}^2$, e cerchiamo la matrice $M_{\mathcal{B}\mathcal{B}'}(f)$.

Avremo:

$f(1, 1, 0) = (2, 0)$; $f(0, 1, 0) = (1, 0)$; $f(0, 1, 1) = (1, 1)$.

Ma i tre vettori $(2, 0)$, $(1, 0)$, $(1, 1)$ non sono le colonne della matrice cercata, perché non esprimono i vettori $f(1, 1, 0)$; $f(0, 1, 0)$; $f(0, 1, 1)$ rispetto alla base $\mathcal{B}'$, ma rispetto alla base canonica (per esempio, si ha $f(0, 1, 1) = (1, 1)_{\mathcal{E}} = 1(1, 0) + 1(0, 1)$)[1]. Per trovare le colonne "giuste" dovremo risolvere:

$$f(1, 1, 0) = (2, 0)_{\mathcal{E}} = a(1, 1) + b(1, -1) = (a + b, a - b)_{\mathcal{B}'}$$

e cioè risolvere il sistema:

$$\begin{cases} a + b = 2 \\ a - b = 0 \end{cases} \to \begin{cases} a + b = 2 \\ -2b = -2 \end{cases} \to \begin{cases} a = 1 \\ b = 1 \end{cases}$$

quindi si ha $f(1, 1, 0) = (2, 0)_{\mathcal{E}} = (1, 1)_{\mathcal{B}'}$; analogamente si ricava:

$$f(0, 1, 0) = (1, 0)_{\mathcal{E}} = (\tfrac{1}{2}, \tfrac{1}{2})_{\mathcal{B}'},$$

$$f(0, 1, 1) = (1, 1)_{\mathcal{E}} = (1, 0)_{\mathcal{B}'}.$$

La matrice cercata è:

$$M_{\mathcal{B}\mathcal{B}'}(f) = \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 1 & \frac{1}{2} & 0 \end{pmatrix}.$$

Esempio 5.4.3. Sia $f : \mathbb{R} \to \mathbb{R}^4$, data da: $f(x) = (2x, 3x, -x, 0)$. Determiniamo la matrice associata alla f rispetto alle basi canoniche; poiché in $\mathbb{R}$ la base canonica è $\mathcal{E} = \{1\}$ e $f(1) = (2, 3, -1, 0)$, si ha:

$$M_{\mathcal{E}\mathcal{E}_4}(f) = \begin{pmatrix} 2 \\ 3 \\ -1 \\ 0 \end{pmatrix}.$$

Con la seguente proposizione si riassume come il rango della matrice $M_{\mathcal{B}\mathcal{B}'}(f)$ determini $\dim(\mathrm{Im} f)$ e $\dim(\ker f)$.

[1] Potremmo notare che i vettori trovati sono le colonne della matrice:

$$M_{\mathcal{B}\mathcal{E}_2}(f) = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Proposizione 5.4.2. *Sia $f : V \to W$, un' applicazione lineare, e sia $\dim V = n$. Sia $A = M_{\mathcal{B}\mathcal{B}'}(f)$ la matrice associata alla f rispetto alle basi $\mathcal{B} \subset V$ e $\mathcal{B}' \subset W$. Allora avremo:*

- $\dim(\mathrm{Im} f) = r(A)$;
- $\dim(\ker f) = n - r(A)$.

Dimostrazione. Per il primo punto, ricordiamo che se $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, le colonne di A sono date dai vettori $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)$ (espressi tramite $\mathcal{B}'$), e che essi sono dei generatori per $\mathrm{Im} f$. Quindi si ottiene che:

$\dim(\mathrm{Im} f) =$ massimo numero di $f(\mathbf{v}_j)$ indipendenti $=$
$=$ massimo numero di colonne indipendenti di $A = r(A)$.

Una base di $\mathrm{Im} f$ sarà data scegliendo $r(A)$ colonne indipendenti di A.

Per quanto riguarda $\ker f$, potremmo concludere immediatamente usando la Proposizione 5.3.4, ma vediamo il risultato direttamente: si ha che $\ker f$ è dato dai $\mathbf{v} \in V$ tali che $f(\mathbf{v}) = \mathbf{0}_W$, e quindi (quando si scrivano i vettori in coordinate rispetto a $\mathcal{B}$) dalle soluzioni di $A \cdot X = O$, e quindi (dal teorema di Rouché–Capelli): $\dim(\ker f) = n - r(A)$. □

Osservazione 5.4.2. Osserviamo che dalla proposizione precedente segue banalmente:

$$\dim(\ker f) + \dim(\mathrm{Im}\ f) = n - r(A) + r(A) = n,$$

ossia l'enunciato della Proposizione 5.3.4. Si è preferito dimostrare quella proposizione indipendentemente per sottolineare che essa non dipende dall'associare una matrice alla f.

Esempio 5.4.4. Consideriamo l'applicazione $f : \mathbb{R}^3 \to \mathbb{R}^3$ definita da $f(x, y, z) = (y - z, x - y, x - z)$. Determinare $\mathrm{Im} f$, $\ker f$.

Calcoliamo f sui vettori della base canonica:

$$f(1,0,0) = (0,1,1), \qquad f(0,1,0) = (1,-1,0), \qquad f(0,0,1) = (-1,0,-1)$$

quindi la matrice $A = M_{\mathcal{E}_3\mathcal{E}_3}(f)$ che rappresenta la f è :

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}, \text{ che, ridotta con Gauss, diviene: } \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Quindi $r(A) = 2 = \mathrm{Im}\ f$. Una base per $\mathrm{Im} f$ sarà data scegliendo due colonne indipendenti di A, ad esempio: $\{(0,1,1), (1,-1,0)\}$.

Il $\ker f$ è dato dalle soluzioni del sistema $A \cdot X = O$, cioè:

$$\begin{cases} y - z = 0 \\ x - y = 0 \\ x - z = 0 \end{cases};$$

nel sistema solo due equazioni sono indipendenti (infatti $r(A) = 2$) e si ha così:

$$\begin{cases} y - z = 0 \\ x - y = 0 \end{cases} \rightarrow \begin{cases} z = y \\ x = -y \end{cases}$$

In conclusione: $\ker f = \{(-y, y, y)\}$; una base per $\ker f$ sarà data da: $\mathcal{B} = \{(-1, 1, 1)\}$ e $\dim(\ker f) = 1$, come già sapevamo, in quanto $\dim(\ker f) = n - r(A) = 3 - 2 = 1$.

Concludiamo la sezione con un'osservazione che è talvolta utile per riconoscere le applicazioni lineari.

Osservazione 5.4.3. Come abbiamo visto, se abbiamo un'applicazione lineare $f : V \to W$ e V, W hanno dimensione finita, allora la f si rappresenta con una matrice $A = M_{\mathcal{B}\mathcal{B}'}(f)$ se i vettori di V e W vengono scritti tramite le coordinate rispetto a due basi fissate $\mathcal{B} \subset V$, $\mathcal{B}' \subset W$.

Quindi ogni applicazione lineare fra spazi vettoriali di dimensione finita si ricollega ad un'applicazione fra due spazi di tipo $\mathbb{R}^n$ e $\mathbb{R}^m$, definita tramite una matrice $A \in \mathbb{R}^{m,n}$.

Da questo fatto segue il seguente risultato:

Proposizione 5.4.3. *Sia $f : V \to W$ una funzione, e $\dim V = n$, $\dim W = m$. Scrivendo la f tramite le coordinate $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_m)$ rispetto a due basi fissate $\mathcal{B} \subset V$, $\mathcal{B}' \subset W$, avremo che la f si scriverà come:*

$f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$, vale a dire

$$\begin{cases} y_1 = f_1(x_1, \ldots, x_n) \\ y_2 = f_2(x_1, \ldots, x_n) \\ \qquad \vdots \\ y_m = f_m(x_1, \ldots, x_n) \end{cases}.$$

Allora la f è lineare se e solo se $f_1, f_2, \ldots, f_m$ sono polinomi nulli o di primo grado senza termine costante nelle $x_1, \ldots, x_n$.

Dimostrazione. Se la f è lineare, per quanto osservato prima:

$$f(x_1, \ldots, x_n) = \left(M_{\mathcal{B}\mathcal{B}'}(f) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)_{\mathcal{B}'}$$

e svolgendo il prodotto si ottiene un vettore

$$(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))_{\mathcal{B}'}$$

le cui componenti, date dal prodotto di una riga della matrice per la colonna delle x_i, sono polinomi di primo grado senza termine costante.

Se invece la f è una funzione definita, come nelle ipotesi, da

$$f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)),$$

ove ogni $f_i = a_{i1}x_1 + \cdots + a_{in}x_n$, allora si ha che $f = f_A$, ove $A = (a_{ij})$, e quindi la f è lineare (vedi Osservazione 5.2.2). □

5.5 Cambiamenti di base, matrici invertibili

Vediamo che nella rappresentazione di applicazioni lineari tramite matrici il prodotto di matrici corrisponde alla composizione di applicazioni lineari:

Proposizione 5.5.1. *Siano V, W, U spazi vettoriali di dimensioni, rispettivamente, n, m, t. Siano $f : V \to W$, $g : W \to U$, applicazioni lineari. Siano poi $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$, basi di V, W e U.*

$$\begin{array}{ccccc} & f & & g & \\ V & \to & W & \to & U \\ \mathcal{B} & & \mathcal{C} & & \mathcal{D} \end{array}$$

Se $A = M_{\mathcal{BC}}(f)$ e $B = M_{\mathcal{CD}}(g)$, allora:

$$M_{\mathcal{BD}}(g \circ f) = BA.$$

Dimostrazione. Sia $\mathbf{v} \in V$ un vettore e X la sua scrittura come vettore colonna rispetto alla base $\mathcal{B}$; allora avremo che il vettore $f(\mathbf{v})$ si scriverà rispetto alla base $\mathcal{C}$ come:

$$C = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}_{\mathcal{C}} = A \cdot X$$

se poi eseguiamo il prodotto per B, avremo che $B \cdot C$ rappresenta (nelle coordinate in $\mathcal{D}$) il vettore che rappresenta $g(f(\mathbf{v}))$:

$$B \cdot C = B(AX) = (BA)X$$

quindi $\forall \mathbf{v} \in V$, il vettore $(g \circ f)(\mathbf{v}) = g(f(\mathbf{v}))$ è rappresentato da $(BA)X$, cioè

$$M_{\mathcal{BD}}(g \circ f) = BA.$$

□

Nel caso particolare che l'applicazione lineare in questione sia l'identità, ci sono interessanti considerazioni da svolgere. Sia $id_V : V \to V$ l'applicazione identità (che è ovviamente lineare), e sia $\dim V = n$; se fissiamo una base qualsiasi $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ di V, avremo sempre che

$$M_{\mathcal{B}\mathcal{B}}(id_V) = I;$$

si verifica immediatamente considerando

$$id_V(\mathbf{v}_1) = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n, \ldots, id_V(\mathbf{v}_n) = 0\mathbf{v}_1 + \cdots + 1\mathbf{v}_n.$$

Se invece in V abbiamo due diverse basi, come $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ e $\mathcal{B}' = \{\mathbf{z}_1, \ldots, \mathbf{z}_n\}$, allora le matrici $M_{\mathcal{B}\mathcal{B}'}(id_V)$ e $M_{\mathcal{B}'\mathcal{B}}(id_V)$ non saranno $= I$, ma avremo (dalla Proposition 5.5.1):

$$M_{\mathcal{B}\mathcal{B}'}(id_V) \cdot M_{\mathcal{B}'\mathcal{B}}(id_V) = M_{\mathcal{B}'\mathcal{B}'}(id_V \circ id)_V = M_{\mathcal{B}'\mathcal{B}'}(id_V) = I$$

$$M_{\mathcal{B}'\mathcal{B}}(id_V) \cdot M_{\mathcal{B}\mathcal{B}'}(id_V) = M_{\mathcal{B}\mathcal{B}}(id_V \circ id_V) = M_{\mathcal{B}\mathcal{B}}(id_V) = I$$

e quindi le due matrici sono l'una l'inversa dell'altra (rispetto al prodotto), cioè :

$$M_{\mathcal{B}\mathcal{B}'}(id_V) = (M_{\mathcal{B}'\mathcal{B}}(id_V))^{-1}.$$

Osserviamo quale sia l'effetto della moltiplicazione per queste matrici (che le rende dei veri e propri "dizionari" fra la scrittura rispetto a una base e quella rispetto a un'altra). Se X è la scrittura di un vettore $\mathbf{v} \in V$, rispetto alla base $\mathcal{B}$, ad esempio

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}},$$

allora, eseguendo la moltiplicazione:

$$M_{\mathcal{B}\mathcal{B}'}(id_V) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{\mathcal{B}'},$$

otterremo di nuovo il vettore $\mathbf{v}$ (stiamo applicando l'identità), ma scritto rispetto alla base $\mathcal{B}'$; l'effetto della moltiplicazione per $M_{\mathcal{B}\mathcal{B}'}(id_V)$ è quindi quello di riscrivere rispetto alla $\mathcal{B}'$ i vettori di V dati tramite le coordinate in $\mathcal{B}$.

Definizione 5.5.1. Dato uno spazio vettoriale V e due sue basi $\mathcal{B}$ e $\mathcal{B}'$, la matrice $M_{\mathcal{B}\mathcal{B}'}(id_V)$ si chiama *matrice del cambiamento di base* (da $\mathcal{B}$ a $\mathcal{B}'$).

Esempio 5.5.1. Sia $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ la base di $\mathbb{R}^3$ data da

$$\mathbf{v}_1 = (1,1,0), \quad \mathbf{v}_2 = (0,1,1), \quad \mathbf{v}_3 = (2,1,2)$$

e $\mathcal{E}_3$ la base canonica. Determiniamo le matrici $M_{\mathcal{BE}}(id)$ e $M_{\mathcal{EB}}(id)$; per la prima il conto è semplice:

$$\begin{aligned} id(\mathbf{v}_1) &= (1,1,0), \\ id(\mathbf{v}_2) &= (0,1,1), \\ id(\mathbf{v}_3) &= (2,1,2). \end{aligned}$$

e poichè i vettori immagine sono già scritti rispetto alla base canonica, si ha:

$$M_{\mathcal{BE}}(id) = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

La matrice $M_{\mathcal{EB}}(id)$ (inversa della precedente) è determinata dai vettori:

$$\begin{aligned} id(\mathbf{e}_1) &= (1,0,0), \\ id(\mathbf{e}_2) &= (0,1,0), \\ id(\mathbf{e}_3) &= (0,0,1), \end{aligned}$$

ma tali vettori vanno scritti in funzione della base $\mathcal{B}$, quindi le tre terne qui sopra NON sono ancora le colonne della matrice cercata! Per la prima colonna, dovremo determinare a, b, c in modo che:

$$(1,0,0) = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3 =$$

$$= a(1,1,0) + b(0,1,1) + c(2,1,2) = (a+2c, a+b+c, b+2c)$$

da cui si ha il sistema (nelle incognite a, b, c):

$$\begin{cases} a + 2c = 1 \\ a + b + c = 0 \\ b + 2c = 0 \end{cases}$$

che ha come soluzione $a = 1/3, b = -2/3, c = 1/3$, cioè:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}_{\mathcal{E}} = \begin{pmatrix} 1/3 \\ -2/3 \\ 1/3 \end{pmatrix}_{\mathcal{B}}.$$

e quest'ultima è la prima colonna di $M_{\mathcal{EB}}(id)$.

Trovare per esercizio le altre due colonne di $M_{\mathcal{EB}}(id)$.

Notiamo che data una qualsiasi matrice $P \in \mathbb{R}^{n,n}$, se $r(P) = n$, si possono pensare le sue colonne come una base di $\mathbb{R}^n$ (poiché avremo che $r(P) = \dim \mathcal{C}(P)$, lo spazio delle colonne di P, e quindi le colonne di P sono n vettori linearmente indipendenti). Allora, per quanto visto sopra, si ha che:

$$P = M_{\mathcal{DE}}(id)$$

e quindi che P è invertibile, essendo la sua inversa $M_{\mathcal{EP}}(id)$.

Ma come caratterizzare le matrici $A \in \mathbb{R}^{n,n}$ che sono invertibili (cioè per cui esiste A^{-1} tale che $A^{-1} \cdot A = A \cdot A^{-1} = I$)? Vediamo che esse sono proprio le matrici di rango massimo:

Proposizione 5.5.2. *Sia $A \in \mathbb{R}^{n,n}$; allora A è invertibile se e solo se $r(A) = n$.*

Dimostrazione. Se $r(A) = n$, abbiamo notato poche righe sopra che la A sarà invertibile, vedendola come una matrice di un cambiamento di base.

Dimostriamo quindi il viceversa: supponiamo adesso che A sia invertibile e vediamo che deve avere rango $= n$. Consideriamo l'applicazione lineare $f_A : \mathbb{R}^n \to \mathbb{R}^n$, definita da $f_A(X) = A \cdot X$ (scrivendo i vettori di $\mathbb{R}^n$ come colonne), risulterà $A = M_{\mathcal{EE}}(f_A)$. Poichè A è invertibile, possiamo considerare anche $f_{A^{-1}} : \mathbb{R}^n \to \mathbb{R}^n$, e avremo che $f_{A^{-1}}(X) = A^{-1} \cdot X$. Allora:

$$\forall X \in \mathbb{R}^n : \qquad (f_A \circ f_{A^{-1}})(X) = A \cdot (A^{-1} \cdot X) = (A \cdot A^{-1}) \cdot X = I \cdot X = X$$

quindi la $f_{A^{-1}}$ è la funzione inversa della f_A; ma per essere invertibile, la f_A deve essere biunivoca, e ciò implica che $r(A) = \dim(\mathrm{Im}\ f_A) = n$. □

Abbiamo visto che l'associazione di una matrice a un'applicazione lineare dipende dalla scelta delle basi; possiamo porci allora la seguente questione: che relazione c'è fra le diverse matrici associate alla stessa applicazione, ma rispetto a basi diverse?

Esempio 5.5.2. Sia $f : \mathbb{R}^2 \to \mathbb{R}^3$, con
$f(x, y) = (x + y, 3x - y, x - y)$.
In genere per studiare un'applicazione come questa, le si associa una matrice usando le basi canoniche nei due spazi; si calcola quindi: $f(1, 0) = (1, 3, 1)$ e $f(0, 1) = (1, -1, -1)$. Da questi si ha:

$$M_{\mathcal{E}_2,\mathcal{E}_3}(f) = \begin{pmatrix} 1 & 1 \\ 3 & -1 \\ 1 & -1 \end{pmatrix}.$$

Consideriamo adesso le basi $\mathcal{B}, \mathcal{B}'$ in $\mathbb{R}^2$ e $\mathbb{R}^3$, rispettivamente:

$$\mathcal{B} = \{\mathbf{v_1}, \mathbf{v_2}\} = \{(1, 1), (1, -1)\};$$

$$\mathcal{B}' = \{\mathbf{w_1}, \mathbf{w_2}, \mathbf{w_3}\} = \{(2, 2, 0), (0, 4, 2), (0, 0, 1)\};$$

(verificare per esercizio che sono effettivamente delle basi). Determiniamo $M_{\mathcal{B},\mathcal{B}'}(f)$; si ha:

$$f(1,1) = (2,2,0) = 1(2,2,0) + 0(0,4,2) + 0(1,0,0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}_{\mathcal{B}'};$$

$$f(1,-1) = (0,4,2) = 0(2,2,0) + 1(0,4,2) + 0(1,0,0) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}_{\mathcal{B}'}.$$

Quindi si ottiene:

$$M_{\mathcal{B},\mathcal{B}'}(f) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Come si può notare, quest'ultima matrice è molto più semplice da usare della precedente; l'esempio mostra che talvolta può essere conveniente usare basi più complicate di quella canonica se la matrice che si ottiene per rappresentare l'applicazione lineare è molto più semplice e facile da usare anche per eventuali computi successivi.

Vediamo allora come sono legate le matrici associate alla stessa applicazione lineare. Sia $f : V \to W$ un'applicazione lineare, e siano $M_{\mathcal{B},\mathcal{B}'}(f)$ e $M_{\mathcal{C},\mathcal{C}'}(f)$ le matrici a essa associate rispetto alle basi $\mathcal{B}, \mathcal{C}$ di V e $\mathcal{B}', \mathcal{C}'$ di W.

Per trovare un legame fra queste due matrici, consideriamo le applicazioni:

$$\begin{array}{ccccccc} & \overset{id_V}{} & & \overset{f}{} & & \overset{id_W}{} & \\ V & \to & V & \to & W & \to & W \\ \mathcal{C} & & \mathcal{B} & & \mathcal{B}' & & \mathcal{C}' \end{array}$$

Si ha che la matrice associata alla composizione $id_W \circ f \circ id_V$ rispetto alle basi sopra indicate è il prodotto delle matrici associate alle singole applicazioni, come abbiamo visto all'inizio della sezione; avremo quindi, poiché $id_W \circ f \circ id_V = f$, che

$$M_{\mathcal{C},\mathcal{C}'}(f) = M_{\mathcal{C},\mathcal{C}'}(id_W \circ f \circ id_V) = M_{\mathcal{B}',\mathcal{C}'}(id_W) \cdot M_{\mathcal{B},\mathcal{B}'}(f) \cdot M_{\mathcal{C},\mathcal{B}}(id_V).$$

Questa relazione fornisce il legame cercato; la $M_{\mathcal{C},\mathcal{C}'}(f)$ si ottiene dalla $M_{\mathcal{B},\mathcal{B}'}(f)$ moltiplicandola a destra e a sinistra per le opportune matrici dei cambiamenti di base (che, se $\dim V = n$ e $\dim W = m$, risulteranno rispettivamente $n \times n$ e $m \times m$).

In particolare, quando $f : V \to V$ è un endomorfismo, e $\mathcal{C} = \mathcal{C}'$, $\mathcal{B} = \mathcal{B}'$, allora si ottiene:

$$M_{\mathcal{C},\mathcal{C}}(f) = M_{\mathcal{B},\mathcal{C}}(id_V) \cdot M_{\mathcal{B},\mathcal{B}}(f) \cdot M_{\mathcal{C},\mathcal{B}}(id_V).$$

Si ha allora che le due matrici dei cambiamenti di base sono una l'inversa dell'altra, cioè, detta P la matrice del cambiamento di base, si ha:

$$M_{\mathcal{C},\mathcal{C}}(f) = P \cdot M_{\mathcal{B},\mathcal{B}}(f) \cdot P^{-1}.$$

Questa situazione è quella descritta dalla seguente definizione:

Definizione 5.5.2. Due matrici $A, B \in \mathbb{R}^{n,n}$ si dicono *simili*, se esiste una matrice invertibile $P \in \mathbb{R}^{n,n}$, tale che: $B = PAP^{-1}$.

Quindi per quanto abbiamo notato in precedenza, matrici associate allo stesso endomorfismo, ma rispetto a una base diversa, sono simili.

5.6 Esercizi

Esercizio 5.6.1. Dire se le seguenti applicazioni siano lineari o no.

1. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 + 2x$.
2. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \sin^2 x + \cos^2 x - 1$.
3. $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x + y, x - y)$.
4. $f : \mathbb{R}^2 \to \mathbb{R}^3$, $f(x, y) = (x, 2x - y, 3xy)$.
5. $f : \mathbb{R} \to \mathbb{R}^3$, $f(x) = (x, x^2, x^3)$.
6. $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x + y + z$.
7. $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x + y + z + 1$.

Esercizio 5.6.2. Determinare la matrice $M_{\mathcal{B}\mathcal{B}'}(f)$ nei seguenti casi:

1. $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (x - y, x + 2z, y - z)$, $\mathcal{B} = \mathcal{B}' = \mathcal{E}_3$.
2. $f : \mathbb{R} \to \mathbb{R}^3$, $f(x) = (2x, 5x, 6x)$, $\mathcal{B} = \{1\}$, $\mathcal{B}' = \mathcal{E}_3$.
3. $f : \mathbb{R}^2 \to \mathbb{R}^3$, $f(x, y) = (x + y, x, y)$, $\mathcal{B} = \{(1, 1), (1, 2)\}$, $\mathcal{B}' = \mathcal{E}_3$.
4. $f : \mathbb{R}^2 \to \mathbb{R}^7$, $f(x, y) = (x, 2x, 3y, x + y, 2x - 4y, x - 12y, 21y)$, $\mathcal{B} = \mathcal{E}_2$, $\mathcal{B}' = \mathcal{E}_7$.
5. Provare a scegliere basi diverse negli esercizi precedenti e trovare la $M_{\mathcal{B}\mathcal{B}'}(f)$ per le basi scelte.

Esercizio 5.6.3. Sia data un'applicazione $f : \mathbb{R}^3 \to \mathbb{R}^3$ tale che

$$f((0, 1, 1)) = (3, 0, 0)$$

$$f((2, 0, 1)) = (2, 1, 2)$$

$$f((2, 1, 2)) = (0, -1, 1).$$

Dimostrare che f non è lineare.

Esercizio 5.6.4. * Si consideri l'endomorfismo di $\mathbb{R}^3$, ove sia fissata la base canonica $\mathcal{E}_3 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, $f : \mathbb{R}^3 \to \mathbb{R}^3$, tale che

$$\begin{cases} f(\mathbf{e}_1) = 2e_1 + e_2 - 3e_3 \\ f(\mathbf{e}_2) = e_1 + 3e_2 - 2e_3 \\ f(\mathbf{e}_3) = f(\mathbf{e}_1) - 2f(\mathbf{e}_2) \end{cases}.$$

Si scriva la matrice associata a f rispetto a $\mathcal{E}_3$.

La f è iniettiva? È suriettiva?

Si considerino in $\mathbb{R}^3$ i tre vettori

$\mathbf{v}_1 = (0,1,2)$, $\mathbf{v}_2 = (3,2,1)$, $\mathbf{v}_3 = (0,1,0)$.

Dopo aver verificato che $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ è una base di $\mathbb{R}^3$, scrivere la matrice del cambiamento di base $M_{\mathcal{E}_3\mathcal{B}}(id_{\mathbb{R}^3})$.

Esercizio 5.6.5. In $\mathbb{R}^3$, fissata la base canonica, si consideri l'endomorfismo definito da

$$\begin{cases} f(\mathbf{e}_1) = m\mathbf{e}_1 + \mathbf{e}_2 \\ f(\mathbf{e}_2) = 4\mathbf{e}_1 + m\mathbf{e}_2 \\ f(\mathbf{e}_3) = -\mathbf{e}_1 - \mathbf{e}_2 + 2\mathbf{e}_3 \end{cases}$$

ove m è un parametro reale. Stabilire, al variare di m, se f sia iniettiva, suriettiva. Dare a m un valore per cui f non sia suriettiva e calcolare, in quel caso, una base per $\ker f$ e $\mathrm{Im} f$.

Esercizio 5.6.6. Siano $\mathcal{B} = (\mathbf{v}_1, \mathbf{v}_2)$ e $\mathcal{B}' = (\mathbf{v}'_1, \mathbf{v}'_2)$ due basi di $\mathbb{R}^2$, ove $\mathbf{v}_1 = (2,1)$, $\mathbf{v}_2 = (0,1)$, $\mathbf{v}'_2 = (1,-1)$, $\mathbf{v}'_2 = (1,1)$. Scrivere la matrice del cambiamento di base $M_{\mathcal{B}\mathcal{B}'}(id_{\mathbb{R}^2})$.

Esercizio 5.6.7. Si consideri l'endomorfismo in $\mathbb{R}^3$, dato da

$f(x,y,z) = (x - z, 2x + y, 2y + z)$.

Si tratta di un isomorfismo? Considerata la base $\mathcal{B}$ costituita dai vettori

$\mathbf{v}_1 = (1,0,1)$, $\mathbf{v}_2 = (2,0,0)$ e $\mathbf{v}_3 = (-1,1,1)$,

determinare la matrice $M_{\mathcal{B}\mathcal{B}}(f)$.

Esercizio 5.6.8. In $\mathbb{R}^2$ si consideri la base $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ con $\mathbf{v}_1 = (1,1)$ e $\mathbf{v}_2 = (-1,0)$, e l'applicazione lineare f che, rispetto alla base $\mathcal{B}$, è associata alla matrice

$$M_{\mathcal{B}\mathcal{B}}(f) = \begin{pmatrix} 3 & -2 \\ 1 & 2 \end{pmatrix}.$$

Determinare la matrice di f rispetto alla base canonica.

Esercizio 5.6.9. In $\mathbb{R}^4$ si consideri l'endomorfismo

$f(x_1, x_2, x_3, x_4) = (y_1, y_2, y_3, y_4)$

definito da:

$$\begin{cases} y_1 = 2x_1 + x_2 + 3x_3 + x_4 \\ y_2 = x_1 - x_2 + 2x_4 \\ y_3 = x_2 + x_3 - x_4 \\ y_4 = 0 \end{cases}.$$

1. Determinare la dimensione e una base per $\mathrm{Im}(f)$.
2. Trovare una base per $Ker(f)$.

Esercizio 5.6.10. * In $\mathbb{R}^4$, in cui si pensa fissata la base canonica $\mathcal{E}_4$, sono dati:

- l'applicazione lineare f tale che

$$\begin{cases} f(\mathbf{e}_1) = \mathbf{e}_1 - 2\mathbf{e}_2 \\ f(\mathbf{e}_2) = \mathbf{e}_3 + \mathbf{e}_4 \\ f(\mathbf{e}_3) = \mathbf{e}_2 \\ f(\mathbf{e}_4) = \mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4 \end{cases}$$

- il sottospazio V di equazioni

$$\begin{cases} x_1 + x_4 = 0 \\ -2x_1 + x_3 - x_4 = 0 \\ x_2 + x_4 = 0 \end{cases}$$

(ove x_1, x_2, x_3, x_4 sono le coordinate in $\mathbb{R}^4$ associate a $\mathcal{E}_4$).

Determinare la dimensione e una base per $f(V)$.

Esercizio 5.6.11. Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare definita da $f(x, y, z) = (hx - y, y, hx + hy + z)$ dove $h \in \mathbb{R}$. Dopo aver determinato i valori di h in corrispondenza dei quali f è un isomorfismo, scrivere per uno di essi un'espressione analitica dell'applicazione lineare inversa f^{-1}.

Esercizio 5.6.12. Si consideri l'endomorfismo $f_{A_h} : \mathbb{R}^3 \to \mathbb{R}^3$ associato alla matrice

$$A_h = \begin{pmatrix} h & 0 & h \\ -1 & 1 & h \\ 0 & 0 & 1 \end{pmatrix}$$

con $h \in \mathrm{R}$. Si stabilisca, al variare di h, se f_{A_h} sia iniettiva, suriettiva.

Esercizio 5.6.13. Sia $f : \mathbb{R}^2 \to \mathrm{R}^3$, ove $f(x, y) = (x + y, x - 2y, x)$. Determinare $M_{\mathcal{B},\mathcal{B}'}(f)$, dove
$\mathcal{B} = \{(1, 1), (0, -1)\}$ e $\mathcal{B}' = \{(1, 1, 1), (1, -2, 0), (0, 0, 1)\}$.

Esercizio 5.6.14. Sia $f : \mathbb{R}^2 \to \mathrm{R}^3$, data da
$f(x, y) = (-x + 3y, 3x - y, 4x)$. Si determini $M_{\mathcal{B},\mathcal{B}'}(f)$ dove

$$\mathcal{B} = \{(1, 1), (1, -1)\} \text{ e } \mathcal{B}' = \{(1, 0, 1), (0, 1, 1), (2, -1, -1)\}.$$

Esercizio 5.6.15. Per ognuna delle seguenti coppie di basi $\mathcal{B}$ e $\mathcal{B}'$ di $\mathbb{R}^2$ determinare $M_{\mathcal{B},\mathcal{B}'}(id_{\mathbb{R}^2})$:

- $\mathcal{B} = \{(1, 0), (0, 1)\}$, $\mathcal{B}' = \{(1, \sqrt{3}), (\sqrt{3}, 1)\}$,
- $\mathcal{B} = \{(1, -1), (9, 1/2)\}$, $B' = \{(1, 0), (2, 1)\}$,

- $\mathcal{B} = \{(2,1),(0,2)\}$, $\mathcal{B}' = \{(\sqrt{5}-1,\sqrt{5}),(1,0)\}$.

Esercizio 5.6.16. Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare data da $f(x,y,z) = (x+y-z, y+z, 2x)$. Determinare la matrice $M_{\mathcal{E},\mathcal{B}}(f)$, dove $\mathcal{E}$ è la base canonica e $\mathcal{B} = \{(1,1,0),(-1,0,1),(1,1,1)\}$.

Esercizio 5.6.17. Sia $f : \mathbb{R}^3 \to \mathbb{R}^4$ l'applicazione lineare $f(x,y,z) = (x+kz, x-y, y-z, x+ky+(1-k)z)$.

Determinare, al variare di $k \in \mathbb{R}$: $\dim \mathrm{Im} f$ e $\dim \ker f$ e dire se la f sia suriettiva e/o iniettiva.

Esercizio 5.6.18. Sia f un endomorfismo di $\mathbb{R}^3$ definito dalla matrice

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

rispetto alla base canonica. Determinare la matrice $M_{\mathcal{B},\mathcal{B}}(f)$ che rappresenta f rispetto alla base $\mathcal{B} = \{(1,1,0),(-1,0,1),(1,1,1)\}$.

Esercizio 5.6.19. Sia f l'endomorfismo di $\mathbb{R}^3$ definito, rispetto alle basi $\mathcal{B}_1 = \{(1,1,1),(1,0,0),(1,0,-1)\}$ e $\mathcal{B}_2 = \{(1,0,0),(2,1,0),(0,-1,2)\}$, dalla seguente matrice:

$$M_{\mathcal{B}_1,\mathcal{B}_2}(f) = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 3 \\ 1 & 0 & 2 \end{pmatrix}.$$

Determinare la matrice $M_{\mathcal{B}_3,\mathcal{B}_4}(f)$ che rappresenta f rispetto alle basi $\mathcal{B}_3 = \{(-1,0,-1),(1,1,1),(1,-1,0)\}$ e $\mathcal{B}_4 = \{(0,0,1),(1,0,1),(2,1,0)\}$.

Esercizio 5.6.20. Sia data l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$, con $f(x,y,z) = (2kx+by-z, (k+1)y+(k-2a)z, (2a-1)z)$.

Dire, al variare di $k \in \mathbb{R}$, quando f sia iniettiva e/o suriettiva. Per i valori di k per cui f non è un isomorfismo, trovare delle basi per $\mathrm{Im}(f)$ e $\ker(f)$.

Esercizio 5.6.21. Sia f un endomorfismo di $\mathbb{R}^3$ tale che $f(x,y,z) = (2x, z-y, y+x)$.

1. Trovare dimensioni e basi di $\mathrm{Im}(f)$ e $\ker(f)$;
2. Determinare dimensioni e basi per $f(U) \cap f(V)$ e $f(U)+f(V)$ dove $U =< (0,0,1),(2,1,0) >$ e $V =< (1,0,1),(1,1,0) >$.

Esercizio 5.6.22. * Sia $f : \mathrm{R}[x] \to \mathrm{R}[x]$ l'endomorfismo rappresentato dalla funzione "derivata"; cioè:

$$f(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0) =$$
$$= n a_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2 x + a_1,$$

vale a dire che per ogni polinomio $p(x) \in \mathbb{R}[x]$, $f(p(x)) = p'(x)$.

Dimostrare che f è lineare.

Determinare $\mathrm{Im} f$, $\ker f$ e una loro base.

Esercizio 5.6.23. * Sia $f : \mathbb{R}[x]_3 \to \mathbb{R}[x]_3$ (ove $\mathbb{R}[x]_3$ è lo spazio dei polinomi in x di grado ≤ 3, più il polinomio nullo) l'endomorfinsmo rappresentato dalla funzione "derivata" (vedi esercizio precedente). Data la base $\mathcal{B} = \{x^3, x^2, x, 1\}$ di $\mathbb{R}[x]_3$, determinare la matrice $M_{\mathcal{B},\mathcal{B}}(f)$.

Esercizio 5.6.24. Sia $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ la base canonica di $\mathbb{R}^4$. Si consideri l'applicazione lineare $f : \mathbb{R}^4 \to \mathbb{R}^4$ definita da

$$f(\mathbf{e}_1) = \mathbf{e}_1, \quad f(\mathbf{e}_2) = \mathbf{0}, \quad f(\mathbf{e}_3) = \mathbf{e}_2, \quad f(\mathbf{e}_4) = \mathbf{e}_3.$$

1. Scrivere la matrice associata a f rispetto alla base canonica di $\mathbb{R}^4$.
2. Determinare la dimensione e una base sia di $\ker f$ sia di $\mathrm{Im} f$.
3. Determinare la dimensione e una base di $f(W)$ dove W è il sottospazio vettoriale di $\mathbb{R}^4$ definito da

$$W = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + x_2 + x_3 = x_2 + x_4 = 0\}.$$

Esercizio 5.6.25. Sia $\mathbb{R}[x]_2$ lo spazio vettoriale dei polinomi a coefficienti reali di grado minore o uguale a 2.

1. Verificare che esiste un unico endomorfismo $f : \mathbb{R}[x]_2 \to \mathbb{R}[x]_2$ tale che

$$f(2+3x) = 3+3x, \ f(x+2x^2) = 1+3x+2x^2, \ f(1-x^2) = -x-x^2.$$

2. Calcolare dimensioni di $\mathrm{Im} f$ e di $\ker f$.

Esercizio 5.6.26. Sia $f : \mathbb{R}^2 \to \mathbb{R}^7$ l'applicazione lineare tale che $f(x, y) = (x, x, y, y, x + y, x - y, -x - y)$.
La f è iniettiva? È suriettiva?
Trovare una base per $\ker f$ e una base per $\mathrm{Im} f$.

Esercizio 5.6.27. Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare tale che $f(x, y) = (x, x, y)$.

1. Scrivere la matrice $M_{\mathcal{E}_2,\mathcal{B}}(f)$ che è associata a f rispetto alle basi $\mathcal{E}_2 = \{(1, 0), (0, 1)\}$ e $\mathcal{B} = \{(1, 1, 0), (0, 1, 1), (0, 0, 1)\}$ di $\mathbb{R}^2$ e $\mathbb{R}^3$ rispettivamente.
2. Scrivere la matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{E}_3$ e quella da $\mathcal{E}_3$ a $\mathcal{B}$.

Esercizio 5.6.28. Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorfismo la cui matrice associata rispetto alla base canonica è:

$$M = \begin{pmatrix} -1 & 1 & 2 \\ 3 & 3 & 4 \\ 2 & 1 & 1 \end{pmatrix}.$$

- Determinare la matrice M' associata a f rispetto alla base $\mathcal{B} = \{(1,1,-1),(1,0,1),(1,1,0)\}$.
- Determinare dimensioni e basi per $\mathbf{Im}f$ e $\ker f$.

Esercizio 5.6.29. * Siano $S(\mathbb{R}^{2,2})$ e $\mathbb{R}[x]_2$ rispettivamente gli spazi vettoriali delle matrici simmetriche reali 2×2 e i polinomi in una variabile a coefficienti reali di grado minore o uguale a 2. Si consideri l'applicazione lineare $f : S(\mathbb{R}^{2,2}) \to \mathbb{R}[x]_2$ tale che:

$$f\left(\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\right) = 2 + 2x + 4x^2,\ f\left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) = x + x^2,$$

$$f\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right) = 1 - 3x - 6x^2.$$

1. Scrivere la matrice associata a f rispetto alle basi

$$\mathcal{B}_{S(\mathbb{R}^{2,2})} = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

e

$$\mathcal{B}_{\mathbb{R}[x]_2} = \{1, x, x^2\}.$$

2. Trovare basi e dimensioni di $\ker f$ e $\mathbf{Im}f$.

Esercizio 5.6.30. Si consideri l'applicazione lineare $f : \mathbb{R}^{2,2} \to \mathbb{R}^{2,2}$ tale che:

$$f\left(\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 + hx_2 + 3x_3 & -x_1 + 2x_2 - x_3 + hx_4 \\ -x_1 + 2x_2 - x_3 + hx_4 & x_1 + 2x_3 - 2x_4 \end{pmatrix}$$

con $h \in \mathbb{R}$.

1. Stabilire per quali valori di $h \in \mathbb{R}$ si ha $\dim(\ker f) = 2$,
2. assegnato a h uno dei valori appena trovati, determinare basi per $\mathbf{Im}f$ e $\ker f$.

Esercizio 5.6.31. Si considerino le seguenti applicazioni e si determini se siano lineari o meno; di quelle lineari si dica se sono isomorfismi:

- $f : \mathbb{R}^{2,2} \to \mathbb{R}^{2,2}$, con

$$f\left(\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}\right) = \begin{pmatrix} 1 & x_1 \\ 2x_2 & x_3 + x_4 \end{pmatrix};$$

- $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 32x$;
- $f : \mathbb{R}^2 \to \mathbb{R}^7$, $f(x,y) = (x,y,x,y,x,y,x+y)$;
- $f : \mathbb{R}^6 \to \mathbb{R}^2$, $f(x_1,x_2,x_3,x_4,x_5,x_6) = (x_1+x_2+x_3, x_4+x_5+x_6)$.

Esercizio 5.6.32. Trovare una applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^2$, tale che $\ker f = \{(x, y, z) | 2x - 3y = 0\}$ e $im f = \{(a, b) | b = 2a\}$.

Esercizio 5.6.33. Dare un esempio di una $f : \mathbb{R}^4 \to \mathbb{R}^3$, applicazione lineare, che abbia $\dim \ker f = 1$. La f trovata è suriettiva? La matrice $M_{\mathcal{E}_4,\mathcal{E}_3}(f)$ ha rango massimo?

Esercizio 5.6.34. ** Sia $f : V \to W$ un'applicazione lineare. Dimostrare che f è iniettiva se e solo se esiste un'applicazione lineare $g : W \to V$ tale che $g \circ f = id_V$.

Esercizio 5.6.35. Siano $f : \mathbb{R}^2 \to \mathbb{R}^2$ e $g : \mathbb{R}^2 \to \mathbb{R}^2$ le due seguenti applicazioni lineari: $f(x, y) = (x - 2y, y)$ e $g(x, y) = (-y, 2x)$. Scrivere le matrici di $f \circ g$ e $g \circ f$ rispetto alla base canonica.

Esercizio 5.6.36. Siano $f : \mathbb{R}^2 \to \mathbb{R}^3$ e $g : \mathbb{R}^3 \to \mathbb{R}^3$ le due seguenti applicazioni lineari: $f(x, y) = (x-y, y, x+y)$ e $g(x, y, z) = (z, 2y-z, x+z)$. Scrivere la matrice rappresentativa $M_{\mathcal{B},\mathcal{E}_3}(g \circ f)$ della funzione $g \circ f$ dove $\mathcal{E}_3$ è la base canonica di $\mathbb{R}^3$ e $\mathcal{B} := \{(1, 1), (0, 1)\}$.

Esercizio 5.6.37. Disegnare l'immagine del vettore $\mathbf{v} = (1, 1)_{\mathcal{E}}$ tramite applicazione lineare: $f : \mathbb{R}^2 \to \mathbb{R}^3$, $f(x, y) = (x + 3y - z, 2y)$. Mostrare graficamente le coordinate di $f(\mathbf{v})$ sia nella base canonica che nella base $\mathcal{B} = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$.

Esercizio 5.6.38. Disegnare il nucleo e l'immagine della seguente applicazione lineare:
$f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (x + 3y - z, y - z, x + 2y)$.

Esercizio 5.6.39. Disegnare il nucleo della seguente applicazione lineare: $f : \mathbb{R}^3 \to \mathbb{R}^2$, $f(x, y, z) = (x + z, x - y - z)$.

Esercizio 5.6.40. Disegnare l'immagine della seguente applicazione lineare: $f : \mathbb{R}^4 \to \mathbb{R}^3$, $f(x, y, z, t) = (x - y, 2x - 2y, 3x - 3y)$.

CAPITOLO 6

Il determinante

Prerequisiti: Generalità sulle matrici; rango di una matrice; eliminazione di Gauss.

Obiettivi: Conoscere le proprietà essenziali del determinante di una matrice quadrata; conoscere il legame fra determinanti e rango di matrici (anche non quadrate).

6.1 Introduzione

Nelle sezioni precedenti abbiamo visto quanto sia importante saper trovare il rango di una matrice: per la risoluzione dei sistemi lineari, per lo studio delle applicazioni lineari e per molte altre questioni. Talvolta quello che abbiamo bisogno di sapere è solo se la matrice che stiamo studiando abbia rango massimo oppure no; a questo scopo esiste uno strumento alternativo a quello dell'eliminazione di Gauss, e cioè il *determinante*, a cui dedichiamo questo capitolo.

In realtà il determinante, da un punto di vista computazionale, non è affatto un metodo efficiente per calcolare il rango di una matrice, lo è molto di più l'eliminazione di Gauss; l'utilità del determinante consiste soprattutto come strumento teorico in molte questioni che coinvolgono il rango di matrici (lo vedremo nel capitolo sugli autovettori).

Solo nel calcolo del rango di matrici di piccola taglia (2×2 o 3×3 ad esempio), che dipendono da paramentri, il determinante può essere un utile metodo alternativo all'eliminazione di Gauss.

Una visualizzazione geometrica del determinante nel caso $n = 3$ si può trovare nel Capitolo 9.

Ci interessa il determinante perché è un'applicazione

$$\det : \mathbb{R}^{n,n} \to \mathbb{R},$$

che si annulla solo sulle matrici che non hanno rango massimo, cioè:

$$\forall A \in \mathbb{R}^{n,n}: \qquad \det A = 0 \iff r(A) \neq n \ .$$

Dovremo mostrare che una tale applicazione esiste, e scoprire come calcolarla.

Cominciamo notando che c'è un caso banale, quello ove $n = 1$. Le matrici 1×1 sono costituite da un solo numero reale: $A = (a_{11})$, per cui il loro rango è 0 se e solo se $a_{11} = 0$, altrimenti $r(A) = 1$; quindi basterà porre $\det A = a_{11}$ e in questo caso avremo la funzione cercata.

Il primo caso significativo è quindi quello $n = 2$. Una generica matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avrà rango minore di due se e solo se le sue righe sono dipendenti, cioè se $\exists k \in \mathbb{R}$, con $(a, b) = k(c, d) = (kc, kd)$.

Assumiamo per il momento che $c \neq 0 \neq d$; allora la condizione precedente si può esprimere come: $k = \frac{a}{c} = \frac{b}{d}$, quindi come $ad = bc$, o ancora:

$$ad - bc = 0. \tag{6.1}$$

Se $c = 0$, $r(A) < 2$ solo se $a = 0$ o $d = 0$ (una riga o una colonna nulla), e anche in quei casi si verifica l'uguaglianza (6.1); similmente se $d = 0$, si ha $r(A) < 2$ per $c = 0$ oppure $b = 0$.

Quindi per $n = 2$ basta porre $\det A = ab - cd$ e, come abbiamo visto, si ha:

$$r(A) < 2 \iff \det A = 0.$$

Questo modo di procedere non ci porta però molto lontano; per le matrici di taglia maggiore abbiamo bisogno di innescare un procedimento di definizione del determinante che possa funzionare per ogni n.

6.2 Il determinante

La definizione del determinante nel caso generico di matrici $n \times n$ è un po' complicata e ci arriveremo per passi successivi, attraverso le due definizioni che seguono.

Definizione 6.2.1. Sia $A \in \mathbb{R}^{n,n}$, $n \geq 2$; $A = (a_{ij})$ e sia $k \leq n$. Allora si dice *minore* (di ordine k) di A ogni sottomatrice quadrata $(k \times k)$ estratta da A eliminando $n - k$ delle sue righe e delle sue colonne.

In particolare indicheremo con A_{ij} il minore $(n-1) \times (n-1)$ ottenuto eliminando in A la i-esima riga e la j-esima colonna.

Collegata alla nozione di minore, abbiamo la seguente:

Definizione 6.2.2. Si dice *complemento algebrico* dell'elemento a_{ij} di A, il numero

$$\mathcal{A}_{ij} = (-1)^{i+j} \det(A_{ij}).$$

Esempio 6.2.1. Sia $A = \begin{pmatrix} 3 & 2 & 1 & 7 \\ 1 & 1 & 0 & 1 \\ 2 & 1 & 3 & 4 \\ 1 & 1 & 0 & 2 \end{pmatrix}$; si avrà $A_{11} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 3 & 4 \\ 1 & 0 & 2 \end{pmatrix}$, $A_{23} = \begin{pmatrix} 3 & 2 & 7 \\ 2 & 1 & 4 \\ 1 & 1 & 2 \end{pmatrix}$; sono invece minori di ordine 2 le matrici: $\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$, oppure $\begin{pmatrix} 3 & 7 \\ 1 & 2 \end{pmatrix}$.
Non siamo invece ancora in grado di calcolare gli $\mathcal{A}_{i,j}$, non avendo ancora definito il determinante per le matrici 3×3; possiamo comunque calcolare i complementi algebrici in una 3×3:

Sia, ad esempio, $A = \begin{pmatrix} 3 & 2 & 7 \\ 1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix}$; allora si ha $A_{12} = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$, e

$$\mathcal{A}_{12} = (-1)^{1+2} \det A_{12} = (-1)(1.0 - 1.2) \; = \; 2.$$

Con la prossima definizione diamo il determinante come funzione che agisce sulle matrici quadrate di qualsiasi grandezza (rimarrà poi da mostrare che il determinante che abbiamo definito ha le proprietà desiderate).

Definizione 6.2.3. *Sia $A \in \mathbb{R}^{n,n}$; $A = (a_{ij})$; allora definiamo nel modo seguente il determinante di A:*

- *se $n = 1$, $\det A = a_{11}$ (l'unico elemento della matrice);*
- *per $n \geq 2$, poniamo:*

$$\det A = a_{11}\mathcal{A}_{11} + a_{12}\mathcal{A}_{12} + \cdots + a_{1n}\mathcal{A}_{1n}.$$

L'espressione di $\det A$ data come sopra si indica in genere come "sviluppo alla Laplace" (rispetto alla prima riga) del determinante di A; in molti testi si possono trovare definizioni alternative di determinante, di cui poi si dimostra che il loro valore è computato dalla formula di Laplace.

A questo punto potremmo avere un dubbio: può sembrare che ci si sia avvitati su noi stessi, utilizzando la nozione di determinante per definire i complementi algebrici, e poi i complementi algebrici per definire il determinante! Ma abbiamo davvero definito qualcosa?

In effetti la Definizione 6.2.3 funziona; in particolare, funziona *per ricorrenza*: noi abbiamo definito il determinante delle matrici 1×1; quindi possiamo (tramite la Definizione 6.2.2) calcolare i complementi algebrici nelle matrici 2×2, e con essi il determinante delle matrici 2×2. Con il determinante delle 2×2 potremo calcolare i complementi algebrici nelle 3×3, e quindi il determinante di quest'ultime, che a sua volta permette di calcolare i complementi nelle 4×4, e così via.

Notiamo anche che la definizione data adesso per il determinante delle 2×2 coincide con quella trovata in precedenza:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}\mathcal{A}_{11} + a_{12}\mathcal{A}_{12} = a_{11}a_{22} - a_{12}a_{21}.$$

La seguente proposizione dice che la funzione "determinante" è tale che per calcolarla possiamo utilizzare una qualsiasi riga o colonna (sviluppo alla Laplace rispetto all i-esima riga o alla j-esima colonna).

Proposizione 6.2.1. *Sia $A \in \mathbb{R}^{n,n}$, $n \geq 2$; $\forall i, j \in \{1, \ldots, n\}$, si ha:*

$$\det A = a_{i1}\mathcal{A}_{i1} + a_{i2}\mathcal{A}_{i2} + \cdots + a_{in}\mathcal{A}_{in} = a_{1j}\mathcal{A}_{1j} + a_{2j}\mathcal{A}_{2j} + \cdots + a_{nj}\mathcal{A}_{nj}.$$

Dimostrazione. Come prima cosa dimostriamo che lo sviluppo del determinante rispetto a una qualsiasi riga è equivalente a quello rispetto alla prima riga che abbiamo usato per definire il determinante stesso.

Procediamo per induzione su $n \geq 2$. Per prima cosa dobbiamo considerare il primo caso, $n = 2$; in questo caso dovremo solo vedere che il $\det A$ si ottiene anche sviluppando alla Laplace rispetto alla seconda riga:

Sia $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$; sviluppando rispetto alla seconda riga si ha:

$$a_{21}\mathcal{A}_{21} + a_{22}\mathcal{A}_{22} = a_{21}(-a_{12}) + a_{22}a_{21} = a_{11}a_{22} - a_{12}a_{21} = \det A.$$

e quindi siamo a posto.

Per $n \geq 3$, supponiamo di sapere che la proprietà vale per le matrici $(n-1) \times (n-1)$ e dimostriamola per le $n \times n$: sia $A \in \mathbb{R}^{n,n}$ e calcoliamo lo sviluppo alla Laplace rispetto alla riga i-esima, con $i \neq 2$;

$$a_{i1}\mathcal{A}_{i1} + a_{i2}\mathcal{A}_{i2} + \cdots + a_{in}\mathcal{A}_{in} =$$

$$= a_{i1}(-1)^{i+1} \det A_{i1} + a_{i2}(-1)^{i+2} \det A_{i2} + \cdots + a_{in}(-1)^{i+n} \det A_{in}$$

ove $A_{i1}, A_{i2}, \ldots, A_{in}$ sono matrici $(n-1)\times(n-1)$; indichiamo con $\mathcal{A}_{ij,hk}$, ove $i \neq h$, $j \neq k$, il numero $(-1)^{i+j} \times (-1)^{h+k} \det A_{ij,hk}$, ove $A_{ij,hk}$ è il minore $(n-2) \times (n-2)$ di A ottenuto cancellando le righe i, h e le colonne j, k; potremo allora scrivere lo sviluppo precedente (sviluppando il determinante dei vari minori rispetto alla loro prima riga):

$$(-1)^{i+1} a_{i1}[a_{12}\mathcal{A}_{i1,12} + a_{13}\mathcal{A}_{i1,13} + \cdots + a_{1n}\mathcal{A}_{i1,1n}] +$$

$$+(-1)^{i+2}\, a_{i2}[a_{11}\mathcal{A}_{i2,11} + a_{13}\mathcal{A}_{i2,13} + \cdots + a_{1n}\mathcal{A}_{i2,1n}] +$$

$$\vdots$$

$$+(-1)^{i+n}\, a_{in}[a_{11}\mathcal{A}_{in,11} + a_{12}\mathcal{A}_{in,12} + \cdots + a_{1n-1}\mathcal{A}_{in,1n-1}]\ .$$

Raccogliendo rispetto ai coefficienti della prima riga di A avremo:

$$a_{11}[(-1)^{i+2}\mathcal{A}_{i2,11} + (-1)^{i+3}\mathcal{A}_{i3,11} + \cdots + (-1)^{i+n}\mathcal{A}_{in,11}] +$$

$$\vdots$$

$$+\ a_{1n}[(-1)^{i+1}\mathcal{A}_{i1,1n} + (-1)^{i+2}\mathcal{A}_{i2,1n} + \cdots + (-1)^{i+n}\mathcal{A}_{in,1n-1}]\ .$$

Ora possiamo dire che la prima espressione fra parentesi quadre è in effetti $\mathcal{A}_{11}$, infatti in essa appare lo sviluppo alla Laplace di A_{11} rispetto alla riga i-esima, e per la nostra ipotesi di induzione, sappiamo che questo corrisponde a $\det A_{11}$. Poiché la stessa cosa vale per tutte le altre espressioni fra parentesi quadre, abbiamo ricavato che

$$a_{i1}\mathcal{A}_{i1} + a_{i2}\mathcal{A}_{i2} + \cdots + a_{in}\mathcal{A}_{in} = a_{11}\mathcal{A}_{11} + a_{12}\mathcal{A}_{12} + \cdots + a_{1n}\mathcal{A}_{1n} = \det A.$$

Che il $\det A$ è anche uguale allo sviluppo alla Laplace rispetto a una qualsiasi colonna, è una conseguenza di $\det A = \det {}^tA$, in quanto nella matrice trasposta si hanno come righe le colonne di A e come minori $({}^tA)_{ij}$ i trasposti ${}^t(A_{ij})$.

Per l'uguaglianza $\det A = \det {}^tA$ rimandiamo alla Proposizione 6.4.2. □

Esempio 6.2.2. Riprendiamo la matrice dell'esempio precedente e calcoliamone il determinante:

$$\det A = \det \begin{pmatrix} 3 & 2 & 7 \\ 1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} =$$

$$= 3(-1)^{1+1} \det \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} + 2(-1)^{1+2} \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} + 7(-1)^{1+3} \det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} =$$

$$= 3(1.0 - 1.2) - 2(1.0 - 1.2) + 7(1.1 - 1.1) = -2.$$

Vediamo che si ottiene lo stesso risultato anche sviluppando il determinante rispetto a una colonna, a esempio la terza:

$$\det A = \det \begin{pmatrix} 3 & 2 & 7 \\ 1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} =$$

$$= 7(-1)^{1+3} \det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + 2(-1)^{2+3} \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} + 0(-1)^{3+3} \det \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} =$$

$$= 7(1.1 - 1.1) - 2(3.1 - 1.2) + 0(3.1 - 1.2) = -2.$$

Esempio 6.2.3. Calcoliamo ora il determinante di una 4×4:

$$\det \begin{pmatrix} 3 & 2 & 0 & 0 \\ -1 & 5 & 0 & 3 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 2 \end{pmatrix} =$$

$$= 3(-1)^{1+1} \det \begin{pmatrix} 5 & 0 & 3 \\ 1 & 0 & 0 \\ 4 & 1 & 2 \end{pmatrix} + 2(-1)^{1+2} \det \begin{pmatrix} -1 & 0 & 3 \\ 2 & 0 & 0 \\ 3 & 1 & 2 \end{pmatrix}$$

abbiamo scelto di sviluppare $\det A$ rispetto alla prima riga perché conteneva due zeri; ora conviene sviluppare i due determinanti 3×3 rispetto alla seconda colonna, che contiene un solo elemento $\neq 0$; otterremo:

$$3(-1)^{2+3} \det \begin{pmatrix} 5 & 3 \\ 1 & 0 \end{pmatrix} - 2(-1)^{2+3} \det \begin{pmatrix} -1 & 3 \\ 2 & 0 \end{pmatrix} = 21 \ .$$

Esiste anche un'altra regola per il calcolo del determinante chiamata

regola di Sarrus

ATTENZIONE: essa può essere utilizzata

ESCLUSIVAMENTE PER LE MATRICI 3×3

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21} \ .$$

Un metodo per ricordare tale regola consiste nel riscrivere le prime due colonne di A alla destra della matrice, e poi sommare tutti i prodotti

eseguiti secondo le diagonali "NordOvest-SudEst" e sottrarre quelli lungo le diagonali "NordEst-SudOvest":

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

Esempio 6.2.4. Riprendiamo ancora la matrice precedente e calcoliamone il determinante usando la regola di Sarrus:

$$\det A = \det \begin{pmatrix} 3 & 2 & 7 \\ 1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} \begin{matrix} 3 & 2 \\ 1 & 1 \\ 1 & 1 \end{matrix} =$$

$$= 3.1.0 + 2.2.1 + 7.1.1 - 7.1.1 - 3.2.1 - 2.1.0 = -2.$$

Calcoliamo (con Laplace) il determinante della seguente matrice:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 \\ 0 & 2 & 1 & 0 \end{pmatrix}$$

poiché la seconda riga contiene tre zero, conviene sviluppare rispetto a essa:

$$\det A = 1 \cdot (-1)^{2+2} \det \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

ed usando Sarrus si ottiene:

$$\det A = 1.0.0 + 0.1.0 + 1.2.1 - 1.0.0 - 1.1.1 - 0.0.2 \ = \ 2 - 1 \ = \ 1.$$

Allora $\det A = 1$.

Ricordiamo che la proprietà che ci interessa per il determinante è che esso si annulli solo sulle matrici di rango $< n$; per esempio in questo caso, avendo trovato che $\det A \neq 0$, vogliamo poter dedurre che $r(A) = 4$; e cioè che i vettori

$$\mathbf{v}_1 = (1,0,0,1), \quad \mathbf{v}_2 = (0,1,0,0), \quad \mathbf{v}_3 = (2,1,0,1) \quad \mathbf{v}_4 = (0,2,1,0),$$

sono indipendenti (e quindi formano una base di $\mathbb{R}^4$).

Per arrivare a dimostrare ciò abbiamo bisogno come prima cosa di vedere quali sono le proprietà della funzione determinante che abbiamo definito.

6.3 Proprietà del determinante

Si vedrà qui quali sono le proprietà fondamentali della funzione determinante, e come da esse si possa dedurre che il determinante si annulla solo sulle matrici di rango non massimo.

1. Se A ha due righe uguali, oppure una riga nulla (con tutti i suoi elementi nulli), allora $\det A = 0$.

2. Il determinante è lineare rispetto alla somma di righe, cioè se si ha:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ b_{i1} + c_{i1} & b_{i2} + c_{i2} & b_{i3} + c_{i3} & \cdots & b_{in} + c_{in} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

allora si avrà : $\det A = \det A_1 + \det A_2$, dove

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & b_{i3} & \cdots & b_{in} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix};$$

$$A_2 = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ c_{i1} & c_{i2} & c_{i3} & \cdots & c_{in} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}.$$

3. Il determinante è lineare rispetto al prodotto di una riga per uno scalare, cioè $\forall\, \lambda \in \mathbb{R}$:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda a_{i1} & \lambda a_{i2} & \lambda a_{i3} & \cdots & \lambda a_{in} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} =$$

$$= \lambda \det \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}.$$

4. Se A' è la matrice ottenuta da A scambiando di posto fra loro due righe, allora $\det A' = -\det A$.

5. Per ogni matrice $A \in \mathbb{R}^{n,n}$ *triangolare superiore*, si ha che

$$\det A = a_{11}a_{22}\cdots a_{nn}.$$

   In particolare, $\forall\, n \geq 1$, sia $I_n \in \mathbb{R}^{n,n}$ la matrice identità di ordine n; allora $\det I_n = 1$.

6. $\det A = \det {}^tA$, ove ${}^tA$ è la *trasposta* di A, cioè la matrice che ha come righe le colonne di A; se $A = (a_{ij})$, allora ${}^tA = (a_{ji})$.

   Una conseguenza immediata di 6. è che le proprietà 1.,...,4. valgono anche enunciate per le *colonne* di A (basta passare da A a ${}^tA$, che ha lo stesso determinante).

7. (*Teorema di Binet:*) $\forall\, A, B \in \mathbb{R}^{n,n}$ si ha :

$$\det(A \cdot B) = (\det A)(\det B).$$

Dimostrazione. Vediamo i singoli punti.

1. Se A ha una riga nulla, basta sviluppare il $\det A$ (alla Laplace) rispetto a quella riga, e otterremo ovviamente $\det A = 0$.

   Anche se ci sono due righe uguali vogliamo vedere che il determinante è nullo. Procediamo per induzione su $n \geq 2$. Se $n = 2$ il determinante è del tipo: $\det \begin{pmatrix} a & b \\ a & b \end{pmatrix} = ab - ab = 0$ e siamo a posto. Per $n \geq 3$, supponiamo di sapere che la proprietà vale per le matrici $(n-1) \times (n-1)$ e dimostriamola per le $n \times n$: se $A \in \mathbb{R}^{n,n}$ ha due righe uguali, sviluppiamo il suo determinante alla Laplace rispetto a una riga che non sia una delle due uguali. Avremo che nell'espressione $\det A = a_{i1}\mathcal{A}_{i1} + a_{i2}\mathcal{A}_{i2} + \cdots + a_{in}\mathcal{A}_{in}$ i complementi algebrici $\mathcal{A}_{ij}$, $j = 1, \ldots, n$ sono tutti ottenuti calcolando determinanti di matrici $(n-1) \times (n-1)$ che hanno due righe nulle, quindi sono tutti zero per l'ipotesi fatta in precedenza. Allora anche $\det A = 0$.

2. Basta sviluppare il determinante rispetto alla riga in questione:

$$\det A = (b_{i1} + c_{i1})\mathcal{A}_{i1} + (b_{i2} + c_{i2})\mathcal{A}_{i2} + \cdots + (b_{in} + c_{in})\mathcal{A}_{in} =$$

$$= b_{i1}\mathcal{A}_{i1} + b_{i2}\mathcal{A}_{i2} + \cdots + b_{in}\mathcal{A}_{in} + c_{i1}\mathcal{A}_{i1} + c_{i2}\mathcal{A}_{i2} + \cdots + c_{in}\mathcal{A}_{in} =$$
$$= \det A_1 + \det A_2.$$

3. Come nel caso precedente:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda a_{i1} & \lambda a_{i2} & \lambda a_{i3} & \cdots & \lambda a_{in} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} =$$
$$= \lambda a_{i1}\mathcal{A}_{i1} + \lambda a_{i2}\mathcal{A}_{i2} + \cdots + \lambda a_{in}\mathcal{A}_{in} =$$
$$\lambda(a_{i1}\mathcal{A}_{i1} + a_{i2}\mathcal{A}_{i2} + \cdots + a_{in}\mathcal{A}_{in}) = \lambda \det(a_{ij}).$$

4. La dimostrazione funziona per induzione come nel caso 1. Per $n = 2$, sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ e $A' = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$. Avremo allora che $\det A' = cb - ad = -\det A$ e siamo a posto. Supponiamo poi di sapere che il determinante cambia di segno scambiando due righe in una matrice $(n-1) \times (n-1)$ e vediamo che ciò vale anche per le matrici $n \times n$. Sia $A \in \mathrm{R}^{n,n}$, $n \geq 3$ e A' come prima ottenuta scambiando di posto due righe. Sviluppiamo il determinante di A' rispetto a una riga, diciamo la i-esima, diversa dalle due scambiate; avremo che nell'espressione:

$$\det A' = a_{i1}\mathcal{A}'_{i1} + a_{i2}\mathcal{A}'_{i2} + \cdots + a_{in}\mathcal{A}'_{in}$$

anche i complementi algebrici $\mathcal{A}'_{ij}$ sono ottenuti dagli $\mathcal{A}_{ij}$ scambiando di posto due righe, quindi, per l'ipotesi induttiva fatta, si ha $\mathcal{A}'_{ij} = -\mathcal{A}_{ij}$, $\forall j \in \{1, \ldots, n\}$. Da qui si ha immediatamente che anche $\det A' = -\det A$.

5. Ancora per induzione: l'enunciato è immediato per $n = 1, 2$. Se lo supponiamo vero per $n-1$, dimostriamolo per n. Sia allora $n \geq 3$, si ha (sviluppando $\det A$ rispetto alla prima colonna):

$$\det A = a_{11}\mathcal{A}_{11} + 0.\mathcal{A}_{21} + \cdots + 0.\mathcal{A}_{n1} =$$
$$= a_{11}.(-1)^{(1+1)} \det A_{11} = a_{11}(a_{22} \cdots a_{nn}).$$

in quanto $\det A_{11} = a_{22} \cdots a_{nn}$ per ipotesi induttiva (A_{11} è una matrice di ordine $(n-1) \times (n-1)$ triangolare superiore).

6. Per la dimostrazione di questa proprietà si rimanda alla Proposizione 6.4.2.

7. Non dimostreremo il teorema di Binet.

□

Osservazione 6.3.1. Dalle Proprietà 1, 2 e 3, si ha che il determinante non cambia se a una riga sommiamo una combinazione lineare delle altre.

Osservazione 6.3.2. Le Proprietà 2 e 3 ci dicono che la funzione determinante è *lineare rispetto alle righe della matrice.*

La Proprietà 4 si indica con il fatto che la funzione det è *alternante* rispetto alle righe della matrice.

La Proprietà 5 si chiama anche *proprietà di normalizzazione.*

Dall'Osservazione 6.3.1 e dalla proprietà 4 si ha:

Proposizione 6.3.1. *Sia $A' \in \mathbb{R}^{n,n}$, ottenuta da A tramite eliminazione di Gauss, allora*

$$\det A' = \pm \det A \, .$$

Dimostrazione. Infatti, ogni passo dell'eliminazione di Gauss o è come nell'Osservazione 6.3.1, oppure come nella proprietà 4; nel primo caso il determinante non cambia, nel secondo cambia solo di segno. Si avrà quindi che $\det A' = \det A$ se abbiamo fatto un numero pari di scambi di righe, e $\det A' = -\det A$ se ne abbiamo fatto un numero dispari. □

Da quest'ultima osservazione segue il risultato che volevamo, che lega l'annullarsi del determinante e il rango della matrice:

Proposizione 6.3.2. *Sia $A \in \mathbb{R}^{n,n}$. Allora si ha che:*

$$\det A = 0 \quad \text{se e solo se} \quad r(A) < n.$$

Dimostrazione. Sia $A' = (a'_{ij})$ la matrice ottenuta da A tramite eliminazione di Gauss; allora $r(A) < n$, se e solo se $r(A') < n$, il che avviene se e solo se $a'_{11}a'_{22} \cdots a'_{nn} = 0$ (perché quando $r(A') < n$ almeno una riga deve essere nulla) e ciò è a sua volta equivalente (per la proprietà 6) a $\det A' = 0$, da cui concludiamo in quanto $\det A' = \pm \det A$. □

Notiamo anche che, dalla proprietà 6., si ha che l'eliminazione di Gauss può costituire anche un modo per calcolare il determinante di una matrice A: basterà eseguire l'eliminazione di Gauss per portarla in forma triangolare, A', e se $r(A) = n$, e $p_1, \ldots, p_n$ sono i pivots di A', avremo che $\det A = (-1)^d p_1 p_2 \cdots p_n$, ove d è il numero di scambi di righe che abbiamo effettuato nella eliminazione di Gauss.

In effetti, un qualsiasi programma per computer che calcoli il determinante di matrici è in genere programmato a farlo utilizzando l'eliminazione di Gauss.

Da quanto abbiamo visto possiamo anche dedurre la seguente

Proposizione 6.3.3. $\det : \mathbb{R}^{n,n} \to \mathbb{R}$ *è l'unica funzione tale che:*

1. *È lineare rispetto alle righe della matrice.*
2. *È alternante rispetto alle righe della matrice.*
3. *È normalizzata.*

Dimostrazione. L'Osservazione 6.3.2 ci dice che la funzione det ha le tre proprietà indicate. Viceversa, se una funzione $D : \mathbb{R}^{n,n} \to \mathbb{R}$ ha le tre proprietà indicate, allora, come abbiamo visto nella Proposizione 6.3.1, si ha che per ogni $A \in \mathbb{R}^{n,n}$, $D(A)$ si può calcolare eseguendo l'eliminazione di Gauss e poi il prodotto degli elementi sulla diagonale principale (con segno positivo se si è eseguito un numero pari di scambi di righe, negativo altrimenti), ma ciò ci dà esattamente $\det A$. □

6.4 Determinante e permutazioni

Svilupperemo in questa sezione un altro possibile approccio al determinante, spesso usato per definirlo; questo approccio utilizza la nozione di permutazione:

Definizione 6.4.1. Si dice *permutazione* (su n oggetti) una qualsiasi funzione biunivoca dall'insieme $\{1, 2, \ldots, n\}$ in sé.

Si può pensare a una permutazione $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ come a un riordinamento della n-upla ordinata $(1, 2, \ldots, n)$, scrivendola come $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. Ad esempio le permutazioni possibili sull'insieme $\{1, 2, 3\}$ sono:

$$(1,2,3);\ (1,3,2);\ (2,1,3);\ (2,3,1);\ (3,1,2);\ (3,2,1).$$

Ove, se consideriamo la permutazione σ definita da $(1, 3, 2)$, intendiamo che $\sigma(1) = 1, \sigma(2) = 3$ e $\sigma(3) = 2$.

In generale, l'insieme S_n delle permutazioni di $\{1, 2, \ldots, n\}$ ha $n!$ elementi, ove

$$n! = n(n-1)(n-2)\ldots 3 \cdot 2 \cdot 1$$

infatti si hanno n scelte per $\sigma(1)$, $n-1$ scelte per $\sigma(2)$, e così via.

Rispetto a n, $n!$ è un numero molto grande: $3! = 3 \cdot 2 \cdot 1 = 6$; $4! = 24$, $5! = 120$, $6! = 720$, $\ldots$, $10! = 3628800$.

Una permutazione che ha come effetto solo lo scambiare di posto due elementi dell'insieme si dice uno *scambio* (ad esempio la (1,3,2) è uno

scambio, mentre la (3,1,2) non lo è). Non è difficile vedere che ogni permutazione si può ottenere da scambi successivi, cioè ogni permutazione è generata da un certo numero di scambi.

Definizione 6.4.2. Si parla di *parità* di una permutazione σ dicendo che essa è *pari* se ottenuta da un numero pari di scambi successivi, *dispari* altrimenti. Si pone $(-1)^\sigma = 1$ se σ è pari, $(-1)^\sigma = -1$ se σ è dispari.

Osservazione 6.4.1. Notiamo che è possibile che una permutazione σ si possa rappresentare in più modi come prodotto di scambi successivi, ma se il numero di scambi in una tale rappresentazione è d, gli altri modi avranno la stessa parità di d, quindi la Definizione 6.4.2 è ben posta. Non dimostreremo qui questa affermazione.

Possiamo ora vedere come si possano utilizzare le permutazioni per ottenere il determinante di una matrice; si ha infatti

Proposizione 6.4.1. *Sia $A = (a_{ij}) \in \mathbb{R}^{n,n}$, allora si ha:*

$$\det A = \sum_{\sigma \in S_n} (-1)^\sigma a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

In vari testi si trova questa come la definizione di determinante (oppure viene usata l'unicità data dalla Proposizione 6.3.3 per definirlo).

Dimostrazione. Basterà dimostrare che la funzione $D : \mathbb{R}^{n,n} \to \mathbb{R}$ definita da

$$D(A) = \sum_{\sigma \in S_n} (-1)^\sigma a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

è lineare e alternante rispetto alle righe di A ed è normalizzata, in quanto allora, utilizzando la Proposizione 6.3.3, avremo che $D(A) = \det A$.

Il fatto che D sia normalizzata si vede considerando che se A è triangolare superiore, ogni prodotto del tipo $a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$ contiene almeno un fattore nullo (un $a_{i\sigma(i)}$ che sta sotto alla diagonale principale), eccetto al più quello ove σ è l'identità; quindi $D(A) = a_{11} a_{22} \cdots a_{nn}$.

L'alternanza si vede notando che se ottengo da A la matrice A' con lo scambiare due righe, ad esempio la i_1-esima e la i_2-esima, ciò ha come conseguenza, nella formula della D, di applicare lo scambio fra i_1 e i_2 a ogni permutazione σ, e quindi cambiarne la parità aggiungendo uno scambio a quelli di cui σ è composta. L'effetto, nella formula della D, sarà quello di cambiare il segno a ogni suo addendo, ottenendo che $D(A') = -D(A)$.

La linearità rispetto a somma fra righe e moltiplicazione di una riga per uno scalare si ricava direttamente dalla formula della D e la lasciamo per esercizio. □

Possiamo usare questa descrizione alternativa del determinante per dimostrare la proprietà 6 del determinante:

Proposizione 6.4.2. *Sia $A \in \mathbb{R}^{n,n}$, allora* $\det A = \det {}^t A$.

Dimostrazione. Usando l'espressione del determinante data dalla proposizione precedente, non è difficile notare che gli addendi che appaiono sono gli stessi per A e per ${}^t A$, infatti basta considerare che ogni prodotto $a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}$ si può riscrivere (a parte l'ordine dei fattori) come $a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(n)n}$ e avremo in questo modo l'uguaglianza fra gli addendi nell'espressione di $\det A$ e quelli nell'espressione di $\det {}^t A$, che (in quanto la parità di σ e σ^{-1} è la stessa) può scriversi:

$$\begin{aligned}\det {}^t A &= \textstyle\sum_{\sigma\in S_n}(-1)^\sigma a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n} = \\ &= \textstyle\sum_{\sigma\in S_n}(-1)^\sigma a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(n)n}.\end{aligned}$$

□

6.5 Applicazioni del determinante

6.5.1 Matrici che dipendono da parametri

Nel discutere il rango di matrici ove appaiono dei parametri (per es. per applicare il teorema di Rouché–Capelli), può essere più semplice calcolare il determinante piuttosto di cercare di calcolare il rango tramite l'eliminazione di Gauss, perché la presenza di paramentri rende più complessa tale operazione. Magari può essere utile calcolare il determinante dopo aver applicato qualche passo dell'eliminazione, per semplificare i conti.

Esempio 6.5.1. Calcolare il rango di $A = \begin{pmatrix} k & 1 & 2 \\ 0 & k & 0 \\ 1 & 3 & 1 \end{pmatrix}$. Calcolando immediatamente il determinante rispetto alla seconda riga, che contiene due zeri:

$$\det \begin{pmatrix} k & 1 & 2 \\ 0 & k & 0 \\ 1 & 3 & 1 \end{pmatrix} = k \det \begin{pmatrix} k & 2 \\ 1 & 1 \end{pmatrix} = k(k-2).$$

Quindi $\det A = k(k-2)$, e il rango sarà massimo (cioè $r(A) = 3$), per $k \neq 0, 2$. Invece se $k = 0$ o $k = 2$, il rango non è massimo; per $k = 0$ o $k = 2$, avremo che $r(A) = 2$, infatti è immediato constatare che le ultime due colonne sono indipendenti (per qualsiasi valore di k).

Esempio 6.5.2. Calcolare il rango di $A = \begin{pmatrix} 1 & 2 & k \\ 1 & k & 0 \\ 1 & 4 & 3 \end{pmatrix}$. Possiamo eseguire il primo passo dell'eliminazione di Gauss, ottenendo:

$$\begin{pmatrix} 1 & 2 & k \\ 0 & k-2 & -k \\ 0 & 2 & 3-k \end{pmatrix}.$$

Possiamo adesso calcolare il determinante rispetto alla prima colonna, ottenendo:

$$\det A = 1 . \det \begin{pmatrix} k-2 & 2 \\ -k & 3-k \end{pmatrix} = -k^2 + 7k - 6.$$

Si ha $k^2 - 7k + 6 = 0$ per $k = 1,\ 6$; allora per $k \neq 1,\ 6$ si ha che $r(A) = 3$, mentre per $k = 1$ o $k = 6$ si ha $r(A) = 2$ (ad esempio le prime due righe sono sempre linearmente indipendenti avendo come pivot 1 e $k - 2$, o $-k$ se $k = 2$).

6.5.2 Matrici non quadrate

Come abbiamo visto, il determinante è definito SOLO per matrici QUADRATE. Ciononostante, calcolandolo sui minori *minori* (vedi Definizione 6.2.1) di un qualunque matrice, anche non quadrata, esso può essere utilizzato per calcolarne il rango. Ad esempio, se $A \in \mathbb{R}^{m,n}$ e un suo minore M di ordine k ha determinante diverso da zero, allora le k righe (o colonne) di M sono linearmente indipendenti, e quindi lo sono anche le k righe (o colonne) corrispondenti in A. Quindi si ha immediatamente che $r(A) \geq k$.

Più precisamente, si dimostra:

Proposizione 6.5.1. *Sia $A \in \mathbb{R}^{m,n}$, $A \neq O$; allora il rango di A è pari a r se e solo se r è l'ordine del massimo minore di A con determinante non nullo.*

Dimostrazione. Per fissare le idee supporremo $m \leq n$ (se vale il viceversa basterà considerare ${}^tA$).

Sia $r(A) = r$; allora ovviamente tutti i minori di ordine $> r$ hanno determinante nullo (non possono avere più di r righe indipendenti perché nella matrice A non ce ne sono). Dobbiamo dimostrare che esiste un minore M di ordine r in A, con $\det M \neq 0$.

Nella matrice A possiamo scegliere r righe indipendenti; prendiamo la sottomatrice A' di A fatta da quelle r righe; allora $A' \in \mathbb{R}^{r,n}$ e $r(A') = r$, perciò ci saranno anche r colonne di A' linearmente indipendenti (il rango per righe è lo stesso del rango per colonne): scegliendo r colonne

indipendenti di A' avremo un minore M di ordine r con determinante $\neq 0$.

Sia invece ora r l'ordine massimo per cui c'è un minore M con determinante $\neq 0$ in A; allora per quanto visto sopra $r(A) \geq r$; ma se fosse $r(A) = r' > r$, allora per il punto precedente esisterebbe un minore M' di ordine r' con $\det M' \neq 0$, assurdo per la nostra ipotesi su r. Quindi $r(A) = r$. □

Esempio 6.5.3. Sia

$$A = \begin{pmatrix} 1 & 0 & 1 & 3 \\ -1 & 2 & 1 & 0 \\ 0 & 2 & 2 & 3 \end{pmatrix}.$$

Il minore 2×2 in alto a sinistra ha determinante diverso da zero, quindi $r(A) \geq 2$. Verificare per esercizio che effettivamente $r(A) = 2$.

Il calcolo dei determinanti dei minori per individuare il rango di una matrice può essere una operazione lunga, ma la si può abbreviare grazie all'*orlatura* dei minori.

Definizione 6.5.1. Sia M un minore $r \times r$ della matrice $A \in \mathbb{R}^{m,n}$, $r < m \leq n$; si dice che un minore M', $(r+1) \times (r+1)$, è ottenuto *orlando* M se M' si ottiene considerando un'altra riga e un'altra colonna oltre a quelle che formano M.

Esempio 6.5.4. Nella matrice

$$A = \begin{pmatrix} 1 & 0 & 1 & 3 \\ -1 & 2 & 1 & 0 \\ 0 & 2 & 2 & 3 \end{pmatrix},$$

i minori $M_1 = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 1 \\ 0 & 2 & 2 \end{pmatrix}$ e $M_2 = \begin{pmatrix} 1 & 0 & 3 \\ -1 & 2 & 0 \\ 0 & 2 & 3 \end{pmatrix}$ sono ottenuti orlando il minore $M = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$.

L'utilizzo dei minori orlati è descritto dalla seguente proposizione (Teorema di Kronecker):

Proposizione 6.5.2. *Sia M un minore $r \times r$ della matrice $A \in \mathbb{R}^{m,n}$, $r < m \leq n$ e si abbia $\det M \neq 0$. Allora $r(A) = r$ se e solo se ogni minore $(r+1) \times (r+1)$ ottenuto orlando M ha determinante nullo.*

Dimostrazione. Poiché $r(M) = r$, se sappiamo che $r(A) = r$, allora i minori $(r+1) \times (r+1)$ che orlano M hanno determinante nullo in quanto tutti i minori $(r+1) \times (r+1)$ di A hanno determinante nullo (per la Proposizione precedente).

Se invece sappiamo che tutti i minori $(r+1) \times (r+1)$ che orlano M hanno determinante nullo si può concludere che $r(A) = r$, perché se si avesse $r(A) \geq r+1$, allora dovrebbero esistere una riga e una colonna indipendenti da quelle che formano M, e scegliendole per formare un minore $(r+1) \times (r+1)$ orlato a M esso avrebbe determinante $\neq 0$, contro l'ipotesi. □

Esempio 6.5.5. Per calcolare il rango della matrice nell' Esempio 6.5.3, invece di calcolare tutti i determinanti dei minori 3×3 (che sono 4), basterà calcolare quello dei minori M_1 e M_2 visti nell' Esempio 6.5.4, in quanto sono gli orlati del minore 2×2 in alto a sinistra che si era osservato avere determinante $\neq 0$.

Esempio 6.5.6. Per calcolare, al variare di k, il rango della matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & k^2-1 \\ 1 & 1 & k & 0 \\ 1 & 2 & 1 & 0 \end{pmatrix}$$

si può osservare che il minore 2×2, M in alto a sinistra, ha determinante $\neq 0$. Quindi $r(A) \geq 2$; per vedere quando $r(A) = 3$, basterà considerare i minori 3×3 orlati di M, e cioè:

$$M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & k \\ 1 & 2 & 1 \end{pmatrix}, \qquad M_2 \begin{pmatrix} 0 & 1 & k^2-1 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \end{pmatrix}.$$

Avremo (sviluppando sulla prima riga); $\det M_1 = k-1$, mentre (sviluppando sull'ultima colonna) si ha $\det M_2 = k^2 - 1$. Si ha allora che per $k = 1$ i due minori hanno rango 2, e così anche $r(A) = 2$, mentre per $k \neq 1$, $r(A) = 3$.

Notiamo che $\det M_2 = 0$ anche per $k = -1$, ma per tale valore $\det M_1 \neq 0$, quindi $r(A) = 3$.

6.5.3 Calcolo dell'inversa

Abbiamo visto che una matrice quadrata $A \in \mathbb{R}^{n,n}$ può ammettere *inversa*, cioè può esistere una matrice A^{-1}, tale che $A \cdot A^{-1} = A^{-1} \cdot A = I$, ove I è la matrice identica. Abbiamo anche visto che una matrice ammette inversa se e solo se $r(A) = n$, cioè se $\det A \neq 0$. Esiste una formula diretta per il calcolo degli elementi della matrice inversa:

Proposizione 6.5.3. *Sia $A \in \mathbb{R}^{n,n}$, $n \geq 2$, tale che $\det A \neq 0$, e sia $B = (b_{ij}) \in \mathbb{R}^{n,n}$, ove:*

$$\forall\, i, j \in \{1, \ldots, n\}, \qquad b_{ij} = \frac{\mathcal{A}_{ji}}{\det A};$$

allora: $A^{-1} = B$.

Dimostrazione. La dimostrazione della validità della formula per la determinazione degli elementi b_{ij} dell'inversa consiste semplicemente nel computare il prodotto $A \cdot B = C$, e verificare che si ottiene $C = I_n$. Se consideriamo gli elementi sulla diagonale di C, avremo che l'i-esimo elemento sarà dato dal prodotto della i-esima riga di A per l'i-esima colonna di B:

$$c_{ii} = a_{i1}b_{1i} + a_{i2}b_{2i} + \cdots + a_{in}b_{ni} =$$

$$= \frac{1}{\det A}(a_{i1}\mathcal{A}_{i1} + a_{i2}\mathcal{A}_{i2} + \cdots + a_{in}\mathcal{A}_{in}) = \frac{\det A}{\det A} = 1.$$

Avremo invece che gli elementi c_{ij}, con $i \neq j$, saranno:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \frac{1}{\det A}(a_{i1}\mathcal{A}_{j1} + a_{i2}\mathcal{A}_{j2} + \cdots + a_{in}\mathcal{A}_{jn}).$$

Ma l'espressione $a_{i1}\mathcal{A}_{j1} + a_{i2}\mathcal{A}_{j2} + \cdots + a_{in}\mathcal{A}_{jn}$ è nulla, in quanto corrisponde al determinante di una matrice ove si sia sostituito alla j-esima riga una copia della i-esima, ottenendo una matrice con due righe uguali, il cui determinante è quindi nullo per la proprietà 1.

Si è quindi visto che C ha tutti termini nulli tranne quelli sulla diagonale, che valgono 1; quindi $C = I_n$. □

Esempio 6.5.7. Sia $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}$, vogliamo determinare A^{-1}.

Possiamo calcolare $\det A$ con Sarrus, ottenendo $\det A = 1$, poi:

$$b_{11} = \mathcal{A}_{11} = (-1)^{1+1}(1 \cdot 1 - 1 \cdot 0) = 1,$$
$$b_{12} = \mathcal{A}_{21} = (-1)(1 \cdot 1 - 2 \cdot 0) = -1,$$
$$b_{13} = \mathcal{A}_{31} = (1 \cdot 0 - 1 \cdot 0) = 0,$$
$$b_{21} = \mathcal{A}_{12} = (-1)(0 \cdot 1 - 2 \cdot 0) = 0,$$
$$b_{22} = \mathcal{A}_{22} = (1 \cdot 1 - 2 \cdot 0) = 1,$$
$$b_{23} = \mathcal{A}_{32} = (-1)(1 \cdot 0 - 0 \cdot 0) = 0,$$
$$b_{31} = \mathcal{A}_{13} = (0 \cdot 1 - 2 \cdot 1) = -2,$$
$$b_{32} = \mathcal{A}_{23} = (-1)(1 \cdot 1 - 2 \cdot 1) = 1,$$
$$b_{33} = \mathcal{A}_{33} = (-1)(1 \cdot 1 - 0 \cdot 1) = 1,$$

e si ottiene che $A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -2 & 1 & 1 \end{pmatrix}$.

6.5.4 Regola di Cramer

Consideriamo i sistemi lineari della forma $A \cdot X = B$, ove $A \in \mathbb{R}^{n,n}$, cioè si abbiano n equazioni e n incognite. Dal teorema di Rouché–Capelli, sappiamo che il sistema ha una unica soluzione se e solo se si ha che $r(A) = r(A|B) = n$. Supponiamo che sia questo il caso; allora necessariamente $\det A \neq 0$.

In questo caso esiste la matrice inversa di A, A^{-1}, e avremo allora:

$$A^{-1} \cdot A \cdot X = A^{-1} \cdot B \quad \Rightarrow \quad I \cdot X = A^{-1} \cdot B \quad \Rightarrow \quad X = A^{-1} \cdot B.$$

Quindi la soluzione X si ottiene sviluppando il prodotto $A^{-1} \cdot B$, ottenendo una formula, detta *Regola di Cramer*, che permette di scrivere direttamente la soluzione del sistema: i valori delle incognite x_j $(j = 1, \ldots, n)$ sono dati da: $x_j = \frac{\det B_j}{\det A}$, ove la matrice B_j si ottiene sostituendo in A la colonna dei termini noti alla colonna j-esima:

$$B_j = \begin{pmatrix} a_{11} & a_{12} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix}.$$

Come si ottiene l'uguaglianza $x_j = \frac{\det B_j}{\det A}$? Abbiamo: $X = A^{-1} \cdot B$; quindi:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{j1} & c_{j2} & \cdots & c_{jn} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_n \end{pmatrix},$$

dove $A^{-1} = (c_{ij})$, cioè (vedi punto precedente) $c_{ij} = \frac{\mathcal{A}_{ji}}{\det A}$. Nell'uguaglianza si ottiene allora (sviluppando il prodotto $A^{-1} \cdot B$):

$$x_j = (c_{j1}, c_{j2}, \ldots, c_{jn}) \cdot B = c_{j1}b_1 + c_{j2}b_2 + \cdots + c_{jn}b_n =$$

$$= \frac{b_1\mathcal{A}_{1j} + b_2\mathcal{A}_{2j} + \cdots + b_n\mathcal{A}_{nj}}{\det A} = \frac{\det B_j}{\det A}.$$

L'ultima uguaglianza può essere facilmente verificata sviluppando $\det B_j$ rispetto alla j-esima colonna (quella dei b_i).

6.5.5 Prodotto vettoriale

Il determinante può essere utilizzato per calcolare il prodotto vettoriale tra vettori in $\mathcal{V}_O^3$, e, conseguentemente, per calcolare volumi e aree. Per queste applicazioni, rimandiamo alla Sezione 9.4.

6.6 Esercizi

Esercizio 6.6.1. Calcolare il determinante delle seguenti matrici:

$$\begin{pmatrix} 15 & 6 \\ 20 & 8 \end{pmatrix}; \quad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 11 & 15 \end{pmatrix}.$$

Esercizio 6.6.2. Calcolare il determinante delle seguenti matrici:

$$\begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix}; \quad \begin{pmatrix} 3 & 1 & 5 \\ 0 & 0 & 2 \\ 1 & 3 & 1 \end{pmatrix}$$

Esercizio 6.6.3. Calcolare il determinante delle seguenti matrici:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}; \quad \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 2 & 1 & 0 \end{pmatrix}.$$

Esercizio 6.6.4. Calcolare il determinante delle matrici $C = AB$ e $D = BA$ dove $A = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 2 & 1 \end{pmatrix}$.

Esercizio 6.6.5. Calcolare il determinante delle matrici $C = AB$ e $D = BA$ dove $A = \begin{pmatrix} 0 & 2 \\ -1 & 1 \\ 1 & 0 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & -1 \end{pmatrix}$.

Esercizio 6.6.6. Calcolare il determinante delle seguenti matrici:

$$\begin{pmatrix} -2 & -2 & -3 \\ 14 & 2 & 7 \\ 0 & 1 & 10 \end{pmatrix}; \quad \begin{pmatrix} 11 & 0 & 12 \\ 0 & -1 & 32 \\ 0 & 4 & 0 \end{pmatrix}.$$

Esercizio 6.6.7. Utilizzando il determinante, valutare il rango delle seguenti matrici al variare del parametro $k \in \mathbb{R}$.

$$\begin{pmatrix} k^2 & 2k \\ k & k+1 \end{pmatrix}; \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & k & 6 \\ 3 & 6 & k+5 \end{pmatrix}.$$

Esercizio 6.6.8. Utilizzando il determinante, valutare il rango delle seguenti matrici al variare del parametro $k \in \mathbb{R}$.

$$\begin{pmatrix} k & 3 \\ 0 & k^2 \end{pmatrix}; \quad \begin{pmatrix} k & 0 & -1 \\ 2-k & k & 1 \\ 0 & -3 & 1 \end{pmatrix}.$$

Esercizio 6.6.9. Utilizzando il determinante, valutare il rango delle seguenti matrici al variare del parametro $k \in \mathbb{R}$.

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & k & 1 \\ 2 & 6 & 3 \end{pmatrix}; \quad \begin{pmatrix} 1 & 1 & 1 \\ 2 & k & k \\ 2k & 4 & k+2 \end{pmatrix}.$$

Esercizio 6.6.10. Utilizzando il determinante, valutare il rango delle seguenti matrici al variare del parametro $k \in \mathbb{R}$.

$$\begin{pmatrix} k-1 & 2 & 3 \\ 0 & k^2-1 & k+1 \\ 0 & 0 & 1+k \end{pmatrix}; \begin{pmatrix} 3-k & 2 & 3 & 21 \\ 0 & k+4 & 6 & 13 \\ 0 & 0 & k-5 & 1 \\ 0 & 0 & 0 & k^2 \end{pmatrix}.$$

Esercizio 6.6.11. Utilizzando il determinante, valutare il rango delle seguenti matrici al variare del parametro $k \in \mathbb{R}$.

$$\begin{pmatrix} -1 & 0 & 1 \\ 0 & -k & 3 \\ 0 & 0 & 3 \end{pmatrix}; \begin{pmatrix} 2 & k & 3 \\ 0 & -k & 3 \\ 0 & 2k & k-3 \end{pmatrix}.$$

Esercizio 6.6.12. Utilizzando il determinante, valutare il rango delle seguenti matrici al variare del parametro $k \in \mathbb{R}$.

$$\begin{pmatrix} k & 0 & 1 \\ 1 & 3 & 0 \\ k^2 & k & 3k \end{pmatrix}; \begin{pmatrix} 1+k & 0 & 0 & 0 \\ 1 & k-2 & 0 & 0 \\ 2 & 3 & k+3 & 0 \\ 1 & 4 & 1 & k^2 \end{pmatrix}.$$

Esercizio 6.6.13. Stabilire se esistono valori reali di k in per ai quali le terne seguenti sono linearmente dipendenti. In caso affermativo trovare i valori di k per i quali si ha la dipendenza lineare:

1. $(2,1,2k+1)$, $(1,2,2)$, $(0,4,1)$;
2. $(1,1,3,1)$, $(k,k,3k,2)$, $(1,1,0,0)$.

Esercizio 6.6.14. Risolvere i seguenti sistemi utilizzando la regola di Cramer (se possibile):

$$\begin{cases} 5x-2y=3 \\ 2x+3y=5 \end{cases}; \quad \begin{cases} x+y+z=4 \\ x+2y+3z=9 \\ x-y+z=2 \end{cases}.$$

Esercizio 6.6.15. Risolvere i seguenti sistemi utilizzando la regola di Cramer (se possibile):

$$\begin{cases} x-y=2 \\ 3x+5y=0 \end{cases}; \quad \begin{cases} 2x+3y-z=1 \\ x-2y+7z=-1 \\ 2x+2y-z=2 \end{cases}.$$

Esercizio 6.6.16. In $\mathbb{R}^4$ si considerino i seguenti vettori $\mathbf{v_1}=(1,0,2,-2)$, $\mathbf{v_2}=(3,-2,-2,0)$, $\mathbf{v_3}=(2,0,1,5)$. Stabilire, utilizzando il determinante, se sono linearmente dipendenti.

Esercizio 6.6.17. Stabilire il massimo numero di vettori indipendenti tra i seguenti vettori di $\mathbb{R}^{20}$:

$$\mathbf{v_1}=(2,1,3,6,0,0,-2,0,-9,7,11,9,0,0,0,0,-1,23,3,1),$$

$$\mathbf{v}_2 = (0, 10, -2, 1, 2, -3, \mathbf{11}, 10, -2, 4, 1, \mathbf{13}, 10, -2, \mathbf{-1}, 2, -3, 0, 0, 0),$$
$$\mathbf{v}_3 = (\mathbf{-2}, 0, 0, \mathbf{0}, 0, 0, 1, 1, \mathbf{4}, 3, 1, -7, 0, \mathbf{2}, 0, \mathbf{-1}, 2, 1, 1, 0),$$
$$\mathbf{v}_4 = (\mathbf{1}, \mathbf{2}, 0, 0, 0, \mathbf{0}, -2, 0, 0, 3, 1, 4, 0, 1, 0, 0, 1, 0, 2, \mathbf{-1}).$$

Esercizio 6.6.18. Stabilire per quali valori del parametro $k \in \mathbb{R}$ i seguenti vettori di $\mathbb{R}^3$ sono linearmente indipendenti: $\mathbf{v}_1 = (k, 2, 1)$, $\mathbf{v}_2 = (0, 1, k)$, $\mathbf{v}_3 = (\mathbf{1}, \mathbf{3}, 2)$.

Esercizio 6.6.19. Stabilire per quali valori del parametro $k \in \mathbb{R}$ i seguenti vettori di $\mathbb{R}^3$ sono linearmente indipendenti: $\mathbf{v}_1 = (0, 1, 3)$, $\mathbf{v}_2 = (k, 2, -k)$, $\mathbf{v}_3 = (-3, -1, 2k)$.

Esercizio 6.6.20. Stabilire se le seguenti matrici sono invertibili e in caso affermativo calcolarne l'inversa:

$$\begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 2 & 0 \\ -2 & 0 & 4 \\ 7 & 1 & 2 \end{pmatrix}.$$

Esercizio 6.6.21. Stabilire se le seguenti matrici sono invertibili e in caso affermativo calcolarne l'inversa:

$$\begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 4 \\ 3 & 1 & 2 \\ -3 & -2 & 12 \end{pmatrix}.$$

Esercizio 6.6.22. Stabilire per quali valori di $k \in \mathbb{R}$ se le seguenti matrici sono invertibili:

$$\begin{pmatrix} k & 2 \\ 2 & k \end{pmatrix}, \begin{pmatrix} -k & 0 & 0 \\ k & k & 0 \\ 1 & k & 1 \end{pmatrix}.$$

Esercizio 6.6.23. Stabilire per quali valori di $k \in \mathbb{R}$ se le seguenti matrici sono invertibili:

$$\begin{pmatrix} 2 & k \\ k & 8 \end{pmatrix}, \begin{pmatrix} 1 & -k & 2+k \\ 0 & k & 0 \\ 3 & -k & 2 \end{pmatrix}.$$

Esercizio 6.6.24. Stabilire per quali valori di $k \in \mathbb{R}$ se le seguenti matrici sono invertibili:

$$\begin{pmatrix} k+1 & 0 & 0 & 0 \\ 0 & k & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 2 & 1 & 3 \end{pmatrix}; \begin{pmatrix} 0 & 0 & 0 & k \\ k & 0 & 0 & 0 \\ 1 & -3 & 1 & 2 \\ 1 & -1 & 0 & 0 \end{pmatrix}.$$

Esercizio 6.6.25. Stabilire per quali valori di $k \in \mathbb{R}$ se le seguenti matrici sono invertibili:

$$\begin{pmatrix} 1 & 2k-3 & 0 \\ 1 & k+2 & 0 \\ k^2 & 0 & 4-k \end{pmatrix}; \begin{pmatrix} 3 & 2 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & k \end{pmatrix}.$$

Esercizio 6.6.26. Calcolare il rango delle seguenti matrici tramite l'utilizzo dei minori:

$$\begin{pmatrix} 0 & 2 & 1 & 0 & 5 \\ 2 & -2 & 4 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 2 & 0 \\ -2 & 2 & 3 & 10 \\ 1 & 7 & -3 & 1 \end{pmatrix}.$$

Esercizio 6.6.27. Calcolare il rango delle seguenti matrici tramite l'utilizzo dei minori:

$$\begin{pmatrix} 4 & -2 & 3 & 0 & 0 \\ 0 & 0 & 7 & 0 & 0 \\ 0 & 0 & 4 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 0 & 7 & 0 & 1 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 3 & 20 \end{pmatrix}.$$

Esercizio 6.6.28. Mostrare che il prodotto di due matrici quadrate invertibili $n \times n$ è ancora una matrice invertibile.

Esercizio 6.6.29. Siano A, B due matrici $n \times n$ tali che $r(A) < n$. Cosa si può dire su $\det(AB)$?

Esercizio 6.6.30. Si risolva l'equazione matriciale

$$AX + BX = C^2 + AX - BX$$

dove A, B, C e X sono matrici quadrate $n \times n$ assumendo che A e B siano invertibili.

Esercizio 6.6.31. Sia A una matrice invertibile. Mostrare che

$$\det A^{-1} = 1/(\det A).$$

Esercizio 6.6.32. Data la matrice $A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 7 \end{pmatrix}$, dire se esiste, e in caso trovarla, una matrice B tale che $AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Se tale B esiste, è possibile affermare che A è invertibile?

Esercizio 6.6.33. * Risolvere, se possibile, col metodo di Cramer (applicato in modo non standard, ma comunque corretto, separando la variabile libera z) il seguente sistema lineare:

$$\begin{cases} x + y = z - 3 \\ x - 3y = z + 1 \end{cases}.$$

Esercizio 6.6.34. Data la matrice

$$A = \begin{pmatrix} x & y & y \\ y & x & y \\ y & y & x \end{pmatrix},$$

calcolare il valore di $\det A$ e dire per quali valori di (x, y) A sia invertibile.

Esercizio 6.6.35. Sia $A \in \mathbb{R}^{4,4}$ dire, giustificando la risposta, se sia vero che $\det A = \det(-A)$. Porsi lo stesso problema in generale, per $A \in \mathbb{R}^{n,n}$.

Esercizio 6.6.36. Considerare l'applicazione det : $\mathbb{R}^{n,n} \to \mathbb{R}$; ci sono valori di n per cui det sia un'applicazione lineare?

Esercizio 6.6.37. Dimostrare che una matrice 3×3 che contenga sette zeri ha determinante nullo.

Esercizio 6.6.38. Sia $A = \begin{pmatrix} 2 & 2 & k \\ 1 & 2 & 0 \\ 0 & 0 & 3k \end{pmatrix}$.

1. Calcolare il rango di A al variare del parametro reale k.
2. Si determini il valore reale di k tale che la matrice A abbia determinante uguale a uno. Per tale valore di k, si calcoli la matrice inversa di A.

Esercizio 6.6.39. [Per chi conosce già il prodotto vettoriale] Calcolare con la regola del determinante e disegnare nello spazio il prodotto vettoriale dei due seguenti vettori: $\mathbf{v}_1 = (1, 1, 2)$, $\mathbf{v}_2 = (0, 0, 2)$.

Esercizio 6.6.40. [Per chi conosce già il prodotto vettoriale] Calcolare con la regola del determinante e disegnare nello spazio il prodotto vettoriale dei due seguenti vettori: $\mathbf{v}_1 = (1, 0, -1)$, $\mathbf{v}_2 = (1, 1, 2)$.

CAPITOLO 7

Autovalori e autovettori

Prerequisiti: Spazi vettoriali, applicazioni lineari, matrici.

Obiettivi: Saper determinare autovalori e autovettori per un endomorfismo (equivalentemente, per una matrice quadrata); saper determinare se un endomorfismo sia semplice, è esprimibile tramite una matrice diagonale.

7.1 Introduzione

Il problema da cui partiamo per esemplificare l'argomento di questo capitolo è il seguente.

Sia $f : \mathcal{V}_O^2 \to \mathcal{V}_O^2$ un'applicazione lineare, ci chiediamo se ci siano vettori $\vec{\mathbf{v}} \in \mathcal{V}_O^2$, $\vec{\mathbf{v}} \neq \mathbf{0}$, tali che la f non fa cambiare la direzione di $\vec{\mathbf{v}}$ (ciò equivale a dire che $f(\vec{\mathbf{v}})$ è un multiplo di $\vec{\mathbf{v}}$ stesso): vogliamo cioè che valga:

$$f(\vec{\mathbf{v}}) = \lambda \vec{\mathbf{v}}.$$

Ciò comporterà che tutta la retta in $\mathcal{V}_O^2$ individuata da $\vec{\mathbf{v}}$ viene trasformata dalla f in se stessa.

Naturalmente cercheremo vettori $\vec{\mathbf{v}} \neq \mathbf{0}$, in quanto per il vettore nullo è banale che $f(\mathbf{0}) = \mathbf{0}$ sia un multiplo di $\mathbf{0}$.

Per facilitare i conti, vediamo un esempio di questo problema nel piano, utilizzando le coordinate in $\mathbb{R}^2$.

Esempio 7.1.1. Sia $f(x, y) = (2x, x + y)$. Cerchiamo, se esistono, dei vettori $\mathbf{v} = (x, y)$ e dei $\lambda \in \mathbb{R}$ tali che $f(x, y) = \lambda(x, y)$.

È facile constatare che per tutti i vettori $\mathbf{v} = (x, x)$ si ha $f(\mathbf{v}) = 2\mathbf{v}$, infatti $f(x, x) = (2x, 2x) = 2(x, x)$.

Quindi su tutto il sottospazio $V_2 \subset \mathbb{R}^2$, $V_2 = \{(x, y) | x = y\}$, f agisce su ogni vettore portandolo nel suo doppio.

Inoltre, se consideriamo le immagini dei vettori del tipo $(0, y)$, si ha che $f(0, y) = (0, y) = 1(0, y)$; allora su tutto il sottospazio $V_1 \subset \mathbb{R}^2$, $V_1 = \{(x, y) \mid x = 0\}$, f agisce su ogni vettore portandolo in se stesso.

Geometricamente ciò significa che la f, vista come una trasformazione del piano in sé, porta le rette V_1 e V_2 in se stesse (su V_1 opera come l'identità, su V_2 porta ogni vettore nel suo doppio, cosicché $f(V_2) = V_2$).

In questo capitolo ci proponiamo di studiare questa questione su ogni spazio vettoriale, determinando come individuare i vettori $\mathbf{v}$ (ed i valori di λ) per cui si verifichi che $f(\mathbf{v}) = \lambda\mathbf{v}$.

Come esempio per mostrare quale possa essere l'utilità di una tale ricerca, consideriamo l'applicazione $f(x, y) = (x+2y, 3x+2y)$; se vogliamo associarle una matrice usando la base canonica otterremo:

$$M_{\mathcal{E}_2\mathcal{E}_2}(f) = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}.$$

Se però usiamo come base $\mathcal{B} = \{(2, 3), (1, -1)\}$, che è data da vettori che vengono trasformati in loro multipli, in quanto $f(2, 3) = (8, 12) = 4(2, 3)$ e $f(1, -1) = (-1, 1) = (-1)(1, -1)$, avremo (calcolarla per esercizio):

$$M_{\mathcal{B}\mathcal{B}}(f) = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}.$$

E cioè che la matrice rispetto a una base di vettori che vengono trasformati in loro multipli è molto più semplice in quanto risulterà una matrice diagonale.

Ci occuperemo anche di questo problema: determinare quando un endomorfismo possa essere rappresentato tramite una matrice diagonale. Osserviamo che questo può significare una notevole semplificazione computazionale: una generica matrice $n \times n$ dipende da n^2 parametri, una matrice $n \times n$ diagonale dipende da soli n parametri; in termini di complessità di calcolo (e di tempo-macchina per un computer) la semplificazione può essere notevolissima, specialmente trattando matrici molto grandi.

7.2 Autovalori, autovettori, autospazi

Vogliamo considerare la nozione di "vettori che vengono trasformati in loro multipli" uno spazio vettoriale qualsiasi:

Definizione 7.2.1. Sia $f : V \to V$ un'applicazione lineare da uno spazio vettoriale V in sé, cioè un *endomorfismo* di V; diremo che un vettore $\mathbf{v} \in V$ è un *autovettore* per la f se $\mathbf{v} \neq \mathbf{0}_V$ ed esiste $\lambda \in \mathrm{R}$ tale che:

$$f(\mathbf{v}) = \lambda\mathbf{v}.$$

A sua volta λ si dice un *autovalore* per la f, e $\mathbf{v}$ un autovettore *relativo* a λ.

Esempio 7.2.1. Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$, data da $f(x, y) = (x + 2y, 3x)$; si ha allora che $f(1, 1) = (3, 3) = 3(1, 1)$. Quindi $(1, 1)$ è un autovettore per f, relativo all'autovalore $\lambda = 3$.

Notiamo che se non avessimo posto $\mathbf{v} \neq \mathbf{0}_V$ nella Definizione 7.2.1, allora ogni $\lambda \in \mathbb{R}$ sarebbe un autovalore per la f in quanto avremmo che $f(\mathbf{0}_V) = \mathbf{0}_V = \lambda \mathbf{0}_V$, qualsiasi sia $\lambda \in \mathbb{R}$.

Notiamo anche che se la f ha un autovettore allora ne ha infiniti, infatti:

Proposizione 7.2.1. *Se $\mathbf{v}$ è un autovettore per un'endomorfismo f, allora lo è anche ogni suo multiplo $a\mathbf{v}$, $\forall a \in \mathbb{R}$, $a \neq 0$.*

Dimostrazione. Se $f(\mathbf{v}) = \lambda \mathbf{v}$, allora anche

$$f(a\mathbf{v}) = af(\mathbf{v}) = a\lambda\mathbf{v} = \lambda(a\mathbf{v}),$$

quindi se $\mathbf{v}$ è un autovettore, lo è anche ogni suo multiplo non nullo. □

Ad ogni autovalore λ si può associare l'insieme di tutti i suoi autovettori:

Definizione 7.2.2. Sia $\lambda \in \mathbb{R}$ un autovalore per un endomorfismo $f : V \to V$; allora l'insieme $V_\lambda = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda \mathbf{v}\} \subset V$ si dice l'*autospazio relativo a* λ (V_λ è costituito da tutti gli autovettori relativi a λ più il vettore nullo).

Proposizione 7.2.2. *Se λ è un autovalore per $f : V \to V$, allora V_λ è un sottospazio di V.*

Dimostrazione. Infatti, per definizione, $\mathbf{0}_V \in V_\lambda$, e inoltre:
$\forall \mathbf{v}, \mathbf{w} \in V_\lambda$, $\forall a, b \in \mathbb{R}$, si ha:

$$f(a\mathbf{v} + b\mathbf{w}) = af(\mathbf{v}) + bf(\mathbf{w}) = a\lambda\mathbf{v} + b\lambda\mathbf{w} = \lambda(a\mathbf{v} + b\mathbf{w}),$$

quindi $a\mathbf{v} + b\mathbf{w} \in V_\lambda$ e quindi V_λ è un sottospazio di V (per la Proposizione 2.4.1) . □

Esempio 7.2.2. Sia $\mathcal{C}^\infty$ lo spazio vettoriale dato da tutte le funzioni $\mathbb{R} \to \mathbb{R}$, infinitamente derivabili. Sia $D : \mathcal{C}^\infty \to \mathcal{C}^\infty$ l'applicazione lineare "derivata" (cioè $D(g) = g'$). Poiché $\ker D = \mathbb{R}$ (le funzioni costanti), si ha che 0 è un autovalore per D, i cui autovettori sono le funzioni costanti: $D(k) = 0.k = 0$, $\forall\, k \in \mathbb{R}$ e quindi $V_0 = \ker D = \mathbb{R}$.

Vediamo anche che ogni $\lambda \in \mathbb{R}$ è un autovalore per D; infatti si ha:

$$\forall\, \lambda \in \mathbb{R}: \quad D(e^{\lambda x}) = \lambda e^{\lambda x},$$

quindi ogni numero reale λ è un autovalore per D (con autovettore $e^{\lambda x}$ e tutti i suoi multipli); il relativo autospazio sarà

$$V_\lambda = \{g(x) = a.e^{\lambda x} | a \in \mathbb{R}\} \subset \mathcal{C}^\infty.$$

Notiamo adesso che se $\lambda = 0$ è un autovalore, allora il relativo autospazio V_0 non è altro che il $\ker f$, infatti

$$V_0 = \{\mathbf{v} \in V \,|\, f(\mathbf{v}) = 0.\mathbf{v} = \mathbf{0}_V\} = \ker f.$$

Da ciò segue

Proposizione 7.2.3. *Un endomorfismo $f : V \to V$ ha $\lambda = 0$ come autovalore se e solo se f non è un isomorfismo.*

Dimostrazione. Si ha infatti:

f non è un isomorfismo $\Leftrightarrow$ $\ker f \neq \{\mathbf{0}_V\} \Leftrightarrow V_0 \neq \{\mathbf{0}_V\} \Leftrightarrow 0$ è un autovalore. □

Esempio 7.2.3. Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$, con $f(x, y, z) = (x-y-z, 2x+y+z, x)$. Si ha che

$$A = M_{\mathcal{E}_3\mathcal{E}_3}(f) = \begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

con $r(A) = 2$ (è facile vedere che $\det A = 0$, o eseguire l'eliminazione di Gauss). Poiché $r(A) = 2 \neq \dim \mathbb{R}^3$, la f non è suriettiva né, quindi, iniettiva (per la Proposizione 5.3.5, o semplicemente, dal fatto che $\dim \ker f = 3 - \dim(\mathrm{Im} f) = 1$). Allora la f, non essendo un isomorfismo, deve avere 0 come autovalore. È facile infatti vedere che $\ker f = \{(0, y, -y)\}$, e quindi

$$f(0, 1, -1) = (0, 0, 0) = 0(0, 1, -1).$$

Come abbiamo visto nel Capitolo 5, se $\dim V = n$, scegliendo una base $\mathcal{B}$ di V possiamo associare alla f una matrice $A = M_{\mathcal{B},\mathcal{B}}(f) \in \mathbb{R}^{n,n}$, tale che ogni $\mathbf{v}$ può esser rappresentato in coordinate rispetto alla base $\mathcal{B}$ da un vettore colonna $X_\mathcal{B} \in \mathbb{R}^{n,1}$. In questo caso $f(\mathbf{v}) = A \cdot X$ (sempre scrivendo rispetto alla base $\mathcal{B}$).

Allora cercare i possibili autovalori λ della f, equivale a determinare i valori $\lambda \in \mathbb{R}$ per i quali esiste un vettore colonna $X \neq \mathbf{0}$ tale che $A \cdot X = \lambda X$.

Definizione 7.2.3. Useremo il termine *autovalore* anche per le matrici quadrate: diremo che $\lambda \in \mathbb{R}$ è un *autovalore per* $A \in \mathbb{R}^{n,n}$ quando esiste $X \in \mathbb{R}^{n,1}$, $X \neq \mathbf{0}$, tale che

$$A \cdot X = \lambda X,$$

e diremo che un tale X è un *autovettore per* A, relativo a λ.

Vediamo come la conoscenza di autovalori per $A \in \mathbb{R}^{n,n}$ ci dia informazioni anche sugli autovalori delle potenze di A e della sua inversa.

Proposizione 7.2.4. *Sia λ un autovalore per la matrice $A \in \mathbb{R}^{n,n}$. Allora:*

- *$\lambda = 0$ è un autovalore per A se e solo se $r(A) < n$;*
- *$\forall\, k \in \mathbb{N}$, $k \geq 2$, λ^k è un autovalore per A^k;*
- *Se A è invertibile, λ^{-1} è un autovalore per A^{-1}.*

Dimostrazione. Il primo punto è una conseguenza immediata di quanto visto per gli endomorfismi: $f_A : \mathbb{R}^n \to \mathbb{R}^n$ ha come autovalore lo 0 se e solo se non è un isomorfismo, e cioè se e solo se $r(A) < n$.

Proviamo il secondo punto per induzione su k; sia $k = 2$, e sia $X \in \mathbb{R}^{n,1}$ un autovettore per A relativo a λ; abbiamo:

$$A^2X = A(AX) = A(\lambda X) = \lambda(AX) = \lambda(\lambda X) = \lambda^2 X,$$

quindi λ^2 è un autovalore per A^2. Supponiamo ora che valga:
$A^{k-1}X = \lambda^{k-1}X$; allora:

$$A^kX = A(A^{k-1}X) = A(\lambda^{k-1}X) = \lambda^{k-1}(AX) = \lambda^{k-1}(\lambda X) = \lambda^k X.$$

e ciò conclude la dimostrazione del secondo punto.

Per la terza affermazione, se A è invertibile, moltiplichiamo per A^{-1} la relazione $AX = \lambda X$; otterremo:

$$A^{-1}(AX) = A^{-1}(\lambda X) = \lambda A^{-1}X,$$

dalla quale, poiché $A^{-1}AX = IX = X$, si ricava: $X = \lambda A^{-1}X$, da cui (dividendo per λ):

$$\lambda^{-1}X = A^{-1}X,$$

quindi λ^{-1} è un autovalore per A^{-1}.

Notiamo che si può dividere per λ in quanto $\lambda \neq 0$, infatti, poiché la A è invertibile, si ha $r(A) = n$ (vedi la Proposizione 5.5.2) e 0 non può essere un autovalore, per il primo punto. □

7.3 Ricerca degli autovalori

Vogliamo ora vedere come determinare gli autovalori e gli autovettori relativi a un endomorfismo f; ci limiteremo al caso in cui $\dim V < +\infty$, per poter associare a f una matrice.

Sia $\dim V = n$, e $f : V \to V$; sia $A = M_{\mathcal{B},\mathcal{B}}(f)$ la matrice associata alla f rispetto alla base $\mathcal{B}$ di V. Allora, se λ è un autovalore per f, e quindi per A, dovrà aversi: $A \cdot X - \lambda X = \mathbf{0}$, cioè (scrivendo al posto di X il suo prodotto per la matrice identità $I \cdot X$): $A \cdot X - \lambda I \cdot X = \mathbf{0}$, e quindi

$$(A - \lambda I) \cdot X = \mathbf{0}.$$

Gli autovalori di A saranno allora le soluzioni X, diverse dal vettore nullo, del sistema omogeneo: $(A - \lambda I) \cdot X = \mathbf{0}$. Sappiamo che tale sistema ammette sempre soluzione (quella nulla), e in particolare ha ∞^{n-r} soluzioni, ove r è il rango della matrice dei coefficienti del sistema, cioè $r = r(A - \lambda I)$.

Avremo quindi che se $r = n$ esiste un' unica soluzione che è necessariamente quella nulla, che non ci dà nessun autovettore X; mentre se $r < n$, esisteranno infinite soluzioni X del sistema, che saranno autovettori per λ; in conclusione, λ può essere un autovalore per A se e solo se $r(A - \lambda I) < n$, cosicché, riassumendo, abbiamo dimostrato la seguente.

Proposizione 7.3.1. *Sia $f : V \to V$, con $\dim V = n$, sia $\mathcal{B}$ una base di V e sia $A = M_{\mathcal{B},\mathcal{B}}(f)$. Allora $\lambda \in \mathbb{R}$ sarà un autovalore per f (e per A), se e solo se la matrice $A - \lambda I$ non ha rango massimo $(= n)$, quindi se e solo se*

$$\det(A - \lambda I) = 0.$$

Questo fornisce il metodo per calcolare gli autovalori di una matrice A (e di un endomorfismo f):

Data una matrice $A \in \mathbb{R}^{n,n}$, i suoi autovalori sono le soluzioni della equazione (detta anche "equazione caratteristica"):

$$p_A(t) = \det(A - tI) = 0.$$

Definizione 7.3.1. Il polinomio $p_A(t) = \det(A - tI)$ si dice il *polinomio caratteristico* della matrice A, e le sue radici sono gli autovalori di A.

Esempio 7.3.1. Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione: $f(x, y) = (2x + y, -x)$. Vogliamo determinarne autovalori e autovettori.

Applicando f ai vettori della base canonica otteniamo:

$$f(1, 0) = (2, -1); \qquad f(0, 1) = (1, 0).$$

e perciò si ha $A = M_{\mathcal{E},\mathcal{E}}(f) = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$. Per trovare gli autovalori di f (e di A) dobbiamo calcolare:

$$\det(A - tI) = \det\left(\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}\right) = \det\begin{pmatrix} 2-t & 1 \\ -1 & -t \end{pmatrix} = t^2 - 2t + 1.$$

Quindi gli autovalori sono le soluzioni dell'equazione $t^2 - 2t + 1 = 0$, cioè $(t - 1)^2 = 0$. La sola soluzione è per $t = 1$.

Allora il solo autovalore è $\lambda = 1$. Gli autovettori sono le soluzioni del sistema: $(A - 1I) \cdot X = \mathbf{0}$:

$$\begin{pmatrix} 2-1 & 1 \\ -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

cioè

$$\begin{cases} x + y = 0 \\ -x - y = 0 \end{cases} \Longrightarrow X = \begin{pmatrix} x \\ -x \end{pmatrix}.$$

Quindi gli autovalori sono i vettori non nulli dell'autospazio:

$$V_1 = \{(x, -x)\} \subset \mathbb{R}^2.$$

Cosa possiamo dire, in generale, sul polinomio $p_A(t)$?

Sviluppando il determinante rispetto ad esempio alla prima colonna, è facile notare che

$$\det(A - tI) = \det\begin{pmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - t \end{pmatrix} =$$

$$= (-1)^n t^n + \cdots + \det A.$$

cioè che il termine noto di $p_A(t)$ è $\det A$ (viene dai termini dello sviluppo di $\det(A - tI)$ che non contengono t), e che il grado del polinomio è sempre $= n$, con coefficiente di t^n pari a $(-1)^n$.

Da queste considerazioni si ha che, se $A \in \mathbb{R}^{n,n}$, A possiede al massimo n autovalori (in quanto $p_A(t)$ ha grado n). Si ha inoltre:

Proposizione 7.3.2. *Se $A \in \mathbb{R}^{n,n}$, con n dispari, allora A possiede almeno un autovalore.*

Dimostrazione. Dobbiamo vedere che il polinomio caratteristico ha almeno una radice, avendo grado dispari. Notiamo che

$$\lim_{t\to+\infty} p_A(t) = -\infty; \quad \lim_{t\to-\infty} p_A(t) = +\infty$$

poiché la parte di massimo grado t^n ha coefficiente $(-1)^n = -1$ in $p_A(t)$. Quindi $p_A(t)$ avrà segno positivo per $t << 0$ e segno negativo per $t >> 0$; poiché $p_A(t)$ è una funzione continua, ci dovrà essere almeno un valore di t per cui $p_A(t) = 0$. □

Il seguente esempio mostra che (per n pari), esistono endomorfismi privi di autovalori.

Esempio 7.3.2. Sia $f : \mathrm{R}^2 \to \mathrm{R}^2$, con $f(x, y) = (2x + 4y, -2x - 2y)$. Si ha che

$$A = M_{\mathcal{E}_2\mathcal{E}_2}(f) = \begin{pmatrix} 2 & 4 \\ -2 & -2 \end{pmatrix},$$

quindi

$$p_A(t) = \det \begin{pmatrix} 2-t & 4 \\ -2 & -2-t \end{pmatrix} = t^2 - 4 + 8 = t^2 + 4.$$

Il polinomio $t^2 + 4$ non ha radici reali, quindi la f non ha autovalori[1].

Esempio 7.3.3. Vogliamo determinare autovalori e autovettori per la matrice:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Calcoliamo il determinante di $A - tI$, e cioè il polinomio caratteristico di A, e cerchiamone le radici:

$$\det(A - tI) = \det \begin{pmatrix} 1-t & 0 & 0 \\ 0 & -t & 1 \\ 0 & -1 & -t \end{pmatrix} = (1-t)(t^2+1) = 0.$$

Il fattore $t^2 + 1$ non si annulla mai[2], quindi l'unico autovalore è dato da $t = 1$.

L'autospazio V_1 è dato dalle soluzioni del sistema:

$$(A - 1I)X = \mathbf{0}, \Rightarrow \begin{cases} 0 = 0 \\ -y + z = 0 \\ -y - z = 0 \end{cases}$$

cioè $V_1 = \{(x, 0, 0)\}$.

[1] NB: stiamo lavorando nel campo REALE: se lavorassimo su C avremmo gli autovalori $\pm 2i$.

[2] Anche in questo caso ci sono valori *complessi* di t che annullano $p_A(t)$, e cioè $t = \pm i$, ma qui ci limitiamo allo studio di autovalori *reali*.

Esempio 7.3.4. Consideriamo il seguente problema: un'associazione venatoria vuole ripopolare di conigli selvatici una zona che attualmente ne è sprovvista. La zona è divisa in tre aree comunicanti A, B, e C; solo nell'area C sarà consentita la caccia.

La previsione è che ogni anno quadruplicherà la popolazione destinata all'area A, quella dell'area B diverrà sette volte maggiore, mentre quella dell'area C, essendo soggetta a caccia, si dimezzerà.

Poiché i "nuovi conigli" nati durante l'anno si redistribuiscono uniformemente nelle tre zone, vorremmo ottenere a fine anno una distribuzione della popolazione nelle tre aree proporzionale a quella iniziale, al fine di avere un ciclo che si ripeterà ogni anno.

Se ciò fosse possibile vogliamo inoltre sapere quale sarà la situazione dopo un anno se abbiamo immesso in tutto 800 conigli.

Soluzione: Siano x, y e z le quantità di conigli immesse nelle tre zone, si avrà :

$$\begin{matrix} A \\ B \\ C \end{matrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \text{ 1 anno } \rightarrow \begin{pmatrix} 4x \\ 7y \\ z/2 \end{pmatrix} = \begin{pmatrix} x+3x \\ y+6y \\ z/2 \end{pmatrix} \rightarrow \text{ redistribuzione } \rightarrow$$

$$\rightarrow \begin{pmatrix} x+x+2y \\ y+x+2y \\ z/2+x+2y \end{pmatrix} = \begin{pmatrix} 2x+2y \\ x+3y \\ x+2y+z/2 \end{pmatrix}$$

Quindi la trasformazione della popolazione dei conigli dopo un anno è data da un'applicazione lineare: $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, con:

$f(x,y,z) = (2x+2y, x+3y, x+2y+\frac{z}{2})$. Si ha

$$M = M_{\mathcal{E}_3\mathcal{E}_3}(f) = \begin{pmatrix} 2 & 2 & 0 \\ 1 & 3 & 0 \\ 1 & 2 & \frac{1}{2} \end{pmatrix}.$$

Se X è la colonna delle distribuzioni iniziali, noi vorremmo che la distribuzione dopo un anno fosse del tipo λX, proporzionale a X, quindi quello che cerchiamo sono autovalori e autovettori di M. Il polinomio caratteristico (sviluppando il determinante rispetto alla terza colonna) sarà:

$$p_M(t) = \det \begin{pmatrix} 2-t & 2 & 0 \\ 1 & 3-t & 0 \\ 1 & 2 & \frac{1}{2}-t \end{pmatrix} = \left(\frac{1}{2}-t\right) \det \begin{pmatrix} 2-t & 2 \\ 1 & 3-t \end{pmatrix} =$$

$$= \left(\frac{1}{2}-t\right)\left(t^2-5t+4\right).$$

Risolvendo la parte di secondo grado, si vede facilmente che le soluzioni di $p_M(t) = 0$ sono:

$$\lambda_1 = \frac{1}{2}, \quad \lambda_2 = 1, \quad \lambda_3 = 4.$$

Cerchiamo i rispettivi autovettori risolvendo i sistemi $(M - \lambda_i I) \cdot X = O$, per $i = 1, 2, 3$:

- $\underline{\lambda_1 = \frac{1}{2}}$. Si ha:

$$\begin{pmatrix} 2-\frac{1}{2} & 2 & 0 \\ 1 & 3-\frac{1}{2} & 0 \\ 1 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \rightarrow \begin{cases} \frac{3x}{2} + 2y = 0 \\ x + \frac{5y}{2} = 0 \\ x + 2y = 0 \end{cases}$$

$$\rightarrow \begin{cases} x = 0 \\ y = 0 \end{cases}.$$

Quindi l'autospazio è : $V_{\frac{1}{2}} = \{(0, 0, z) | z \in \mathbb{R}\}$.

- $\underline{\lambda_2 = 1}$. Si ha:

$$\begin{pmatrix} 2-1 & 2 & 0 \\ 1 & 3-1 & 0 \\ 1 & 2 & \frac{1}{2}-1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \rightarrow \begin{cases} x + 2y = 0 \\ x + 2y = 0 \\ x + 2y - \frac{z}{2} = 0 \end{cases}$$

$$\rightarrow \begin{cases} x = -2y \\ z = 0 \end{cases}.$$

Quindi l'autospazio è : $V_1 = \{(-2y, y, 0) | y \in \mathbb{R}\}$.

- $\underline{\lambda_1 = 4}$. Si ha:

$$\begin{pmatrix} 2-4 & 2 & 0 \\ 1 & 3-4 & 0 \\ 1 & 2 & \frac{1}{2}-4 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \rightarrow \begin{cases} -2x + 2y = 0 \\ x - y = 0 \\ x + 2y - \frac{7z}{2} = 0 \end{cases}$$

$$\rightarrow \begin{cases} x = y \\ z = \frac{6x}{7} \end{cases}.$$

Quindi l'autospazio è: $V_4 = \{(x, x, \frac{6x}{7}) | x \in \mathbb{R}\}$.

Vediamo ora di interpretare questi risultati alla luce del nostro problema.

L'autovalore $\lambda_1 = \frac{1}{2}$, con gli autovettori $\{(0, 0, z)\}$ ci danno una possibile soluzione al nostro problema: immettere tutti i conigli nella sola area C! Ma questa è una scelta che solo dei cacciatori miopemente ingordi farebbero. Certo, per il primo anno avrebbero a disposizione molta selvaggina cacciabile, ma essa si dimezzerebbe ogni anno, e dopo poco la popolazione diventerebbe irrisoria.

L'autovalore $\lambda_2 = 1$ non dà nessuna soluzione al nostro problema: infatti gli autovettori $\{(-2y, y, 0)\}$ non hanno alcun interesse in quanto non possiamo immettere in una zona una quantità di conigli negativa!

Quindi l'ultima speranza per i nostri cacciatori è l'autovalore $\lambda_3 = 4$, con gli autovettori $\{(x, x, \frac{6x}{7})\}$. Se la quantità totale immessa deve essere di 800 conigli, avremo:

$$x + x + \frac{6x}{7} = 800 \quad \text{e quindi} \quad \frac{20x}{7} = 800 \text{ e } x = 280.$$

Allora questa è la soluzione ottimale per i cacciatori: immetteranno 280 conigli rispettivamente nelle aree A e B, e 240 nell'area C. Dopo un anno la popolazione totale sarà quadruplicata e nella zona C ci saranno ben 960 conigli da cacciare, molti di più dei 400 che avremmo avuto con la soluzione "ingorda" data dal primo autovalore.

Osserviamo ora il polinomio $P_a(t)$ dal punto di vista della sua possibile scomposizione in fattori; come abbiamo visto anche negli esempi precedenti, in generale il polinomio caratteristico di una matrice $A \in \mathbb{R}^{n,n}$ si potrà scrivere come:

$$p_A(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_s)^{n_s}q(t)$$

ove $\lambda_1, \ldots, \lambda_s \in \mathbb{R}$, $s \leq n$, sono le radici reali di $p_A(t)$, mentre il fattore $q(t)$ non ha radici reali.

Definizione 7.3.2. Sia $A \in \mathbb{R}^{n,n}$ e $\lambda_1, \ldots, \lambda_s$, $s \leq n$, siano i suoi autovalori. Sia

$$p_A(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_s)^{n_s}q(t).$$

Il valore $n_i \in \mathbb{N}$ è detto *molteplicità algebrica* dell'autovalore λ_i.

Si dice invece *molteplicità geometrica* di un autovalore λ_i il numero $m_i = \dim V_{\lambda_i}$, la dimensione dell'autospazio relativo a λ_i.

Nell'Esempio 7.3.3, si ha un unico autovalore $\lambda_1 = 1$, con $n_1 = 1$, mentre $q(t) = t^2 + 1$; invece nell' Esempio 7.3.1 si ha ancora il solo autovalore $\lambda_1 = 1$, ma $n_1 = 2$, mentre $q(t) = 1$.

Si ha ovviamente (poiché $p_A(t)$ ha grado n): $n_1 + n_2 + \cdots + n_s \leq n$, e il grado di $q(t)$ è $n - n_1 - \cdots - n_s$, in particolare quando, come nell'Esempio 7.3.1, $n = n_1 + \cdots + n_s$, allora $q(t) = 1$, e più semplicemente non lo scriviamo nello sviluppo di $p_A(t)$.

Si ha invece che $m_i = n - r(A - \lambda_i I)$, in quanto ricordiamo che V_{λ_i} è lo spazio delle soluzioni del sistema $(A - \lambda_i I) \cdot X = \mathbf{0}$.

Si ha sempre (ma non lo dimostreremo qui) che

$$1 \leq m_i \leq n_i.$$

Si osservi che la prima disuguaglianza è ovvia perché se λ_i è un autovalore, $\dim V_{\lambda_i} \geq 1$.

7.4 Endomorfismi semplici, diagonalizzazione

Data un'applicazione lineare $f : V \to V$, dove $\dim V = n$, è di particolare rilevanza determinare se esista una base di V fatta di autovettori per f, in quanto se useremo una tale base (se esiste) per associare a f una matrice, otterremo "la più semplice" matrice che si possa associare alla f. Vediamo in che senso; sia $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base fatta tutta di autovettori per f, e siano $\mu_1, \ldots, \mu_n$ i relativi autovalori (eventualmente anche non distinti). La matrice $A = M_{\mathcal{B},\mathcal{B}}(f)$ si determinerà calcolando:

$$f(\mathbf{v}_1) = \mu_1\mathbf{v}_1 = \mu_1\mathbf{v}_1 + 0.\mathbf{v}_2 + \cdots + 0\mathbf{v}_n = \begin{pmatrix} \mu_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{\mathcal{B}};$$

$$f(\mathbf{v}_2) = \mu_2\mathbf{v}_1 = 0\mathbf{v}_1 + \mu_2.\mathbf{v}_2 + \cdots + 0\mathbf{v}_n = \begin{pmatrix} 0 \\ \mu_2 \\ \vdots \\ 0 \end{pmatrix}_{\mathcal{B}};$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$f(\mathbf{v}_n) = \mu_n\mathbf{v}_n = 0\mathbf{v_1} + 0.\mathbf{v}_2 + \cdots + \mu_n\mathbf{v}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \mu_n \end{pmatrix}_{\mathcal{B}}.$$

Si ha quindi:

$$A = \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mu_n \end{pmatrix} = D_{\mu_1,\ldots,\mu_n},$$

Cioè la matrice A è una matrice diagonale, che ha come elementi della sua diagonale principale proprio gli autovalori.

Definizione 7.4.1. Un endomorfismo $f : V \to V$ si dice *semplice* se esiste una base di V fatta di autovettori per f.

Una matrice $A \in \mathbb{R}^{n,n}$ si dice *diagonalizzabile* se è simile a una matrice diagonale, cioè se esistono una matrice invertibile $P \in \mathbb{R}^{n,n}$ e una matrice diagonale $D \in \mathbb{R}^{n,n}$ tali che $A = P \cdot D \cdot P^{-1}$.

Proposizione 7.4.1. *Sia $f : V \to V$, con $\dim V = n$. Allora f è semplice se e solo se, per ogni base $\mathcal{B}$ di f, la matrice $M_{\mathcal{BB}}(f)$ è diagonalizzabile.*

Dimostrazione. È immediato (per quanto osservato prima) che f è semplice se e solo se è rappresentabile con una matrice diagonale, scegliendo come base una base $\mathcal{C}$ fatta di autovettori. A sua volta ciò equivale a dire che ogni matrice che rappresenta la f (cioè comunque si scelga la base di V) sarà diagonalizzabile, infatti data una qualsiasi base $\mathcal{B}$ di V, la matrice $M_{\mathcal{BB}}(f)$ è legata alla matrice diagonale $D = M_{\mathcal{CC}}(f)$ dalla relazione:

$$M_{\mathcal{BB}}(f) = M_{\mathcal{CB}}(id_V) \cdot M_{\mathcal{CC}}(f) \cdot M_{\mathcal{BC}}(f);$$

che dà la similitudine cercata essendo del tipo:

$$M_{\mathcal{BB}}(f) = P^{-1}DP.$$

□

Esempio 7.4.1. Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$, con $f(x, y, z) = (x + z, y + z, x + y)$. f è semplice? Cerchiamo gli autovalori di f. La matrice associata a f rispetto alla base canonica è : $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, e quindi gli autovalori sono le soluzioni dell'equazione:

$$p_A(t) = \det \begin{pmatrix} 1-t & 0 & 1 \\ 0 & 1-t & 1 \\ 1 & 1 & -t \end{pmatrix} = 0$$

sviluppando il determinante rispetto alla prima colonna si ottiene:

$$(1-t)(t^2 - t - 1) - (1-t) = (1-t)(t^2 - t - 2) = 0$$

che dà come soluzioni: $t = 1, -1, 2$. Determiniamo gli autovettori relativi ai tre autovalori trovati:

- $\underline{\lambda_1 = -1}$: V_{-1} è dato dalle soluzioni del sistema $(A-(-1)I)X = 0$:

$$\begin{cases} 2x + z = 0 \\ 2y + z = 0 \\ x + y + z = 0 \end{cases} \Rightarrow \begin{cases} -2x = z \\ y = x \end{cases}$$

  Quindi $V_{-1} = \{(x, x, -2x)\}$.

- $\underline{\lambda_2 = 1}$: V_1 è dato dalle soluzioni del sistema:

$$\begin{cases} z = 0 \\ z = 0 \\ x + y - z = 0 \end{cases} \Rightarrow \begin{cases} z = 0 \\ y = -x \end{cases}$$

  Quindi $V_1 = \{(x, -x, 0)\}$.

- $\underline{\lambda_3 = 2}$: V_2 è dato dalle soluzioni del sistema:

$$\begin{cases} -x + z = 0 \\ -y + z = 0 \\ x + y - 2z = 0 \end{cases} \Rightarrow \begin{cases} x = z \\ y = x \end{cases}$$

Quindi $V_2 = \{(x, x, x)\}$.

Consideriamo adesso gli autovettori $(1, 1, -2)$, $(1, -1, 0)$, $(1, 1, 1)$, ognuno dei quali è un generatore di uno dei tre autospazi. I tre vettori sono linearmente indipendenti (vederlo per esercizio) e quindi costituiscono una base di R^3, che denotiamo con $\mathcal{B}$. Vediamo ora di rappresentare f tramite questa base:

$$f(1,1,-2) = (-1,-1,2) = -1(1,1,-2)+0(1,-1,0)+0(1,1,1) = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}_{\mathcal{B}},$$

$$f(1,-1,0) = (1,-1,0) = 0(1,1,-2) + 1(1,-1,0) + 0(1,1,1) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}_{\mathcal{B}},$$

$$f(1,1,1) = (2,2,2) = 0(1,1,-2) + 0(1,-1,0) + 2(1,1,1) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}_{\mathcal{B}}.$$

Quindi $M_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = D_{-1,1,2}$, e f è semplice. Inoltre si ha anche che A è diagonalizzabile, infatti avremo (se $\mathcal{E}$ è la base canonica):

$$M_{\mathcal{E}\,\mathcal{E}}(f) = A = P^{-1} D_{-1,1,2} P$$

ove $M_{\mathcal{B},\mathcal{E}}(id) = P^{-1}$, $M_{\mathcal{E},\mathcal{B}}(id) = P$.

Vogliamo adesso un criterio per determinare se un'applicazione lineare f sia semplice (o una matrice $A \in \mathbb{R}^{n,n}$ sia diagonalizzabile), senza dover eseguire tutta la ricerca di una base di autovettori come nell'esempio precedente.

Richiamiamo la seguente definizione (vedi Capitolo 2):

Definizione 7.4.2. Se $V_1, \ldots, V_s$ sono sottospazi di uno spazio vettoriale V, si dice *spazio somma* di essi il sottospazio

$$V_1 + V_2 + \cdots + V_s = \{\mathbf{v} \in V \mid \exists \mathbf{v}_i \in V_i,\ i = 1, \ldots, s,\ \mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_s\}.$$

Proposizione 7.4.2. *Sia $f : V \to V$, $\dim V = n$, e siano $\lambda_1, \dots, \lambda_s$ gli autovalori (distinti) di f. Allora*

$$\dim(V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_s}) = \dim V_{\lambda_1} + \dim V_{\lambda_2} + \cdots + \dim V_{\lambda_s}.$$

Dimostrazione. Sia $\dim V_{\lambda_i} = m_i$, e siano $\mathcal{B}_i = \{\mathbf{v}_{1,i}, \dots, \mathbf{v}_{m_i,i}\}$ basi dei V_{λ_i}, $i = 1, \dots, s$. Per dimostrare la proposizione vogliamo vedere che $\mathcal{B} = \cup_{i=1}^{s} \mathcal{B}_i = \{\mathbf{v}_{1,1}, \dots, \mathbf{v}_{m_1,1}, \dots, \mathbf{v}_{1,s}, \dots, \mathbf{v}_{m_s,s}\}$ è una base di $V_{\lambda_1} + \cdots + V_{\lambda_s}$.

Ovviamente $\mathcal{B}$ è un insieme di generatori per $V_{\lambda_1} + \cdots + V_{\lambda_s}$, infatti se $\mathbf{v} \in V_{\lambda_1} + \cdots + V_{\lambda_s}$, si avrà $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_s$, con $\mathbf{v}_i \in V_{\lambda_i}$; inoltre, poiché ogni $\mathcal{B}_i$ è una base per V_{λ_i}, esisteranno $\alpha_{1,i}, \dots, \alpha_{m_i,i} \in \mathbb{R}$ tali che $\mathbf{v}_i = \alpha_{1,i}\mathbf{v}_{1,i} + \cdots + \alpha_{m_i,i}\mathbf{v}_{m_i,i}$, $i = 1, \cdots, s$, e quindi

$$\mathbf{v} = \alpha_{1,1}\mathbf{v}_{1,1} + \cdots + \alpha_{m_1,1}\mathbf{v}_{m_1,1} + \cdots + \alpha_{1,s}\mathbf{v}_{1,s} + \cdots + \alpha_{m_s,s}\mathbf{v}_{m_s,s}.$$

Se dimostriamo che i $\mathbf{v}_{i,j}$ sono anche linearmente indipendenti avremo finito. Per questo si procede per induzione su s. Per $s = 1$ l'enunciato è ovvio, infatti in questo caso $\mathcal{B} = \mathcal{B}_1$ e i $\mathbf{v}_{1,j}$ sono quindi linearmente indipendenti. Supponiamo adesso che l'enunciato sia vero per $s-1$ (cioè supponiamo che i vettori in $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_{s-1}$ siano indipendenti), e vediamo che allora vale anche per s. Siano $\alpha_{i,j} \in \mathbb{R}$ tali che

$$\alpha_{1,1}\mathbf{v}_{1,1} + \cdots + \alpha_{m_1,1}\mathbf{v}_{m_1,1} + \cdots + \alpha_{1,s}\mathbf{v}_{1,s} + \cdots + \alpha_{m_s,s}\mathbf{v}_{m_s,s} = \mathbf{0}_V.$$

Dobbiamo dimostrare che allora gli $\alpha_{i,j}$ sono tutti nulli. Applicando la f ai due lati di questa uguaglianza avremo:

$$f(\alpha_{1,1}\mathbf{v}_{1,1} + \cdots + \alpha_{m_s,s}\mathbf{v}_{m_s,s}) = f(\mathbf{0}_V)$$

che dà

$$\lambda_1\alpha_{1,1}\mathbf{v}_{1,1} + \cdots + \lambda_s\alpha_{m_s,s}\mathbf{v}_{m_s,s} = \mathbf{0}_V.$$

Moltiplicando la prima uguaglianza per λ_s e sottraendola dall'ultima si ottiene:

$$(\lambda_1 - \lambda_s)\alpha_{1,1}\mathbf{v}_{1,1} + \cdots + (\lambda_{s-1} - \lambda_s)\alpha_{m_{s-1},s-1}\mathbf{v}_{m_{s-1},s-1} = \mathbf{0}_V.$$

Ma, poiché $\lambda_i - \lambda_s \neq 0$, $i = 1, \dots, s-1$, e, per ipotesi induttiva, $\mathbf{v}_{1,1}, \dots, \mathbf{v}_{m_{s-1},s-1}$ sono linearmente indipendenti, avremo allora che $\alpha_{1,1} = \cdots = \alpha_{m_{s-1},s-1} = 0$. Quindi la prima uguaglianza diviene:

$$\alpha_{1,s}\mathbf{v}_{1,s} + \cdots + \alpha_{m_s,s}\mathbf{v}_{m_s,s} = \mathbf{0}_V,$$

il che implica che $\alpha_{1,1} = \cdots = \alpha_{m_s,s} = 0$ (perché i $\mathbf{v}_{1,1}, \dots, \mathbf{v}_{m_s,s}$ sono indipendenti essendo una base di V_{λ_s}). Quind gli $\alpha_{i,j}$ sono tutti nulli. □

Dalla Proposizione 7.4.2 ricaviamo immediatamente il criterio desiderato:

Proposizione 7.4.3. *Sia $f : V \to V$, $\dim V = n$, e siano $\lambda_1, \ldots, \lambda_s$ gli autovalori (distinti) di f. Allora f è semplice se e solo se:*
$m_1 + \cdots + m_s = n$.

Dimostrazione. Supponiamo che $m_1 + \cdots + m_s = n$, allora, per la Proposizione 7.4.2, si ha che che $\dim(V_{\lambda_1} + \cdots + V_{\lambda_s}) = n$, e cioè che $V_{\lambda_1} + \cdots + V_{\lambda_s} = V$ e quindi che $\mathcal{B} = \{\mathbf{v}_{1,1}, \ldots, \mathbf{v}_{m_1,1}, \ldots, \mathbf{v}_{1,s}, \ldots, \mathbf{v}_{m_s,s}\}$ è una base di V, fatta di autovettori di f (ogni $\mathbf{v}_{i,j} \in V_{\lambda_i}$), quindi f è semplice.

Viceversa, se esiste una base di V, $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ fatta di autovettori di f, allora ogni $\mathbf{v}_i$ appartiene necessariamente a $V_{\lambda_1} + \cdots + V_{\lambda_s}$, quindi $V_{\lambda_1} + \cdots + V_{\lambda_s} = V$ (contenendo una base di V) e quindi che $\dim(V_{\lambda_1} + \cdots + V_{\lambda_s}) = m_1 + \cdots + m_s = n$. □

Notiamo che poiché si ha sempre $m_1 + \cdots + m_s \leq n_1 + \cdots + n_s \leq n$, quando f è semplice si ha anche $n_1 + \cdots + n_s = n$. In particolare avremo che $p_A(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_s)^{n_s}$, cioè $p_A(t)$ ha solo soluzioni reali (la parte in $q(t)$ non appare).

Inoltre, quando f è semplice si ha anche $m_i = n_i$, $i = 1, \ldots, s$.

In particolare avremo:

Proposizione 7.4.4. *Se $f : V \to V$, $\dim V = n$, ha n autovalori distinti, $\lambda_1, \ldots, \lambda_n$, allora f è semplice.*

Dimostrazione. La proposizione segue dal fatto che in questo caso abbiamo $n_i = 1$, $\forall\, i = 1, \ldots, n$. Allora si ha anche $m_1 = m_2 = \cdots = m_n = 1$, in quanto $1 \leq m_i \leq n_i$, $\forall\, i = 1, \ldots, n$, e quindi $m_1 + \cdots + m_n = n$. □

Ad esempio, osserviamo che nell'Esempio 7.4.1, una volta trovati gli autovalori $1, -1, 2$, potevamo immediatamente concludere che f era semplice (per la Proposizione 7.4.4, avendo tre autovalori distinti). Naturalmente se avessimo voluto trovare una base di autovettori avremmo dovuto comunque eseguire il procedimento ivi svolto per determinare gli autovettori.

Esempio 7.4.2. Riprendiamo l'esempio 7.3.4; avendo trovato tre autovalori distinti ($\lambda_1 = \frac{1}{2}$, $\lambda_2 = 1$ e $\lambda_3 = 4$), sappiamo che la f che rappresenta il cambiamento annuale della popolazione dei conigli è *semplice*, e una base di autovettori è data ad esempio (prendendo un generatore per ogni autospazio) da:

$$\mathcal{B} = \{\mathbf{v}_1 = (0, 0, 1), \mathbf{v}_2 = (-2, 1, 0), \mathbf{v}_3 = (7, 7, 6)\},$$

cosicché si avrà :

$$M_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} = D_{\frac{1}{2},1,4}.$$

Nella situazione descritta si optava per la scelta di distribuzione degli 800 conigli nelle tre zone a disposizione in modo da sfruttare l'autovalore $\lambda_3 = 4$, e cioè $(A, B, C) = (280, 280, 240) \in V_{\lambda_3}$; tale vettore, è pari a $40\mathbf{v}_3$, quindi scritto rispetto alla base $\mathcal{B}$, risulta: $(0, 0, 40)_{\mathcal{B}}$.

Vogliamo ora vedere cosa accadrà dopo cinque anni, facendo la scelta descritta (assumendo che per ogni anno il cambiamento della popolazione rispetti la stessa legge). La popolazione finale ci sarà data da:

$$D^5_{\frac{1}{2},1,4} \cdot \begin{pmatrix} 0 \\ 0 \\ 40 \end{pmatrix} = \begin{pmatrix} \frac{1}{2^5} & 0 & 0 \\ 0 & 1^5 & 0 \\ 0 & 0 & 4^5 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 40 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1024 \times 40 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 40960 \end{pmatrix}.$$

Una crescita esplosiva, con ben $40.960 \times 800 = 32768000$ conigli, e dopo altri cinque anni salirebbe a più di 33 miliardi!

Naturalmente nella realtà il nostro modello, al crescere della popolazione dei conigli, non rispetterebbe più la situazione effettiva in quanto una grossa popolazione avrebbe più difficoltà a trovare le risorse per aumentare sempre allo stesso tasso di crescita ogni anno; in ogni modo il modello ci fornisce l'indicazione del rischio di un aumento ecocessivo di una popolazione. Una situazione di questo tipo si è avuta in Australia, ove i conigli furono introdotti dall'Europa, nel 1856. Essi trovarono un enorme territorio ricco di risorse a loro favorevoli e privo di predatori e parassiti: la loro riproduzione fu esponenziale con effetti devastanti sulle colture e l'estinzione di intere specie vegetali. Combattuti in molti modi, e più efficacemente con l'introduzione di malattie e virus mirati che ne hanno distrutto anche più del 90%, i ceppi resistenti hanno sempre iniziato nuove moltiplicazioni esplosive.

ATTENZIONE: Quanto stabilisce la Proposizione 7.4.4 NON è una condizione necessaria affinché f sia semplice! Un errore molto comune fra gli studenti, nello stabilire se una data f sia semplice, è quello di trovare gli autovalori, e se non si hanno n autovalori distinti, affermare che perciò la f non è semplice! Quello che bisogna considerare è la somma delle molteplicità geometriche degli autovalori (Proposizione 7.4.3).

Esempio 7.4.3. L'applicazione identità $id_V : V \to V$ è ovviamente semplice, in quanto ogni vettore non nullo è un suo autovettore (in quanto $id_V(\mathbf{v}) = 1\mathbf{v}$), e perciò ogni sua base è una base di autovettori. Qualunque base $\mathcal{B}$ si scelga per V (assumiamo $\dim V = n$), si avrà :

$$M_{\mathcal{B},\mathcal{B}}(id_V) = I_n.$$

Abbiamo che c'è un unico autovalore possibile: $\lambda_1 = 1$ (non n autovalori distinti!), ma di molteplicità geometrica: $m_1 = \dim V_1 = n$.

Esercizio 7.4.1. Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$, con $f(x, y) = (y, kx + y)$. Discutere al variare di k se f sia semplice.

Soluzione: La matrice $A = M_{\mathcal{E}\mathcal{E}}(f)$ è $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ k & 1 \end{pmatrix}$; avremo quindi $p_A(t) = \det(\boldsymbol{A} - t\boldsymbol{I}) = \det\begin{pmatrix} -t & 1 \\ k & 1-t \end{pmatrix} = t^2 - t - k$. Risolvendo l'equazione $t^2 - t - k = 0$ si ha: $t = \frac{1 \pm \sqrt{1+4k}}{2}$.

Per $1 + 4k < 0$, cioè $k < -1/4$, avremo che $p_A(t) = 0$ non ha soluzioni reali, quindi f, non possedendo autovalori reali, non è semplice.

Per $1 + 4k > 0$, cioè $k > -1/4$, avremo due autovalori reali distinti: $\lambda_1 = \frac{1-\sqrt{1+4k}}{2}$ e $\lambda_2 = \frac{1+\sqrt{1+4k}}{2}$, e quindi (per il Corollario 7.4.4), f è semplice.

Quando $k = -1/4$, l'unico autovalore è $\lambda = 1/2$, e $p_A(t) = (t - \frac{1}{2})^2$. Per vedere se la f è semplice dobbiamo determinare $m_\lambda = \dim V_\lambda$ e f sarà semplice se $m_\lambda = n_\lambda = 2$. Si ha $m_\lambda = n - r(\boldsymbol{A} - \lambda\boldsymbol{I})$ e

$\boldsymbol{A} - 1/2\boldsymbol{I} = \begin{pmatrix} -1/2 & -1 \\ -1/4 & 1/2 \end{pmatrix}$, con $r(\begin{pmatrix} -1/2 & -1 \\ -1/4 & 1/2 \end{pmatrix}) = 1$, quindi $m_\lambda = \dim V_\lambda = 2 - 1 = 1$ e f non è semplice.

7.5 Esercizi

Esercizio 7.5.1. Calcolare autovalori e autovettori della seguente applicazione lineare: $f : \mathbb{R}^2 \to \mathbb{R}^2$ dove $f(x, y) = (x - 2y, 3y + 4x)$.

Esercizio 7.5.2. Disegnare autovettori e le loro rispettive immagini dell'esercizio precedente.

Esercizio 7.5.3. Calcolare autovalori e autovettori della seguente applicazione lineare: $f : \mathbb{R}^3 \to \mathbb{R}^3$, dove

$f(x, y, z) = (x - y + z, x - 2y + z, x + 3z)$.

Esercizio 7.5.4. Disegnare autovettori e le loro rispettive immagini dell'esercizio precedente.

Esercizio 7.5.5. Calcolare autovalori e autovettori della seguente applicazione lineare: $f : \mathbb{R}^3 \to \mathbb{R}^3$ dove $f(x, y, z) = (x + z, -y + z, z)$.

Esercizio 7.5.6. Stabilire innanzitutto se la seguente applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ è ben definita e motivarne la risposta:

$$f(1, 1, 0) = (5, 5, 0),$$

$$f(0, 1, 1) = (0, 1, 2),$$

$$f(1, 0, 1) = (-1, 0, 0).$$

Trovarne poi autovalori e autovettori.

Esercizio 7.5.7. Dire se $f : \mathbb{R}^3 \to \mathbb{R}^3$ sia semplice al variare di k, con $f(x, y, z) = (x - y, y + z, x + ky + z)$.

Esercizio 7.5.8. Se $f : \mathbb{R}^4 \to \mathbb{R}^4$, con $f(x, y, z, w) = (w, z, y, x)$, determinarne autovalori e autovettori. La f è semplice?

Esercizio 7.5.9. Determinare autovalori e autovettori della seguente applicazione lineare: $f : \mathbb{R}^2 \to \mathbb{R}^2$ con $f(x, y) = (x, -y)$.

Esercizio 7.5.10. Se $A \in \mathbb{R}^{2,2}$ non ha autovalori, possiamo dire che anche A^2 non ne ha?

Esercizio 7.5.11. Sia $A \in \mathbb{R}^{n,n}$ e λ un autovalore per A. È vero che λ^{-1} è un autovalore per A^{-1}?

Esercizio 7.5.12. In $\mathbb{R}^3$, fissata la base canonica, si consideri l'endomorfismo definito da

$$\begin{cases} f(\underline{e}_1) = m\underline{e}_1 + \underline{e}_2 \\ f(\underline{e}_2) = 2\underline{e}_1 + m\underline{e}_2 \\ f(\underline{e}_3) = -\underline{e}_1 - \underline{e}_2 + 2\underline{e}_3 \end{cases}$$

ove m è un parametro reale. Stabilire, al variare di m, se f sia semplice.

Esercizio 7.5.13. Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$, con $f(x, y, z) = \left(2y, \frac{1}{2}x, 7x + z\right)$. Stabilire se la f sia semplice e trovarne gli autospazi.

Esercizio 7.5.14. Disegnare autovettori, le loro rispettive immagini e gli autospazi dell'esercizio precedente.

Esercizio 7.5.15. Dimostrare che l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ così definita

$$f(x, y, z) = (x, 2x - y + z, x + 5y + 3z)$$

è semplice. Trovare una base di autovettori e scrivere la matrice che rappresenta la f in tale base.

Esercizio 7.5.16. Disegnare autovettori, le loro rispettive immagini e gli autospazi dell'esercizio precedente.

Esercizio 7.5.17. Dimostrare che l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ così definita

$$f(x, y, z) = (x, 2x - x + y + z, -x + 2z)$$

è semplice. Trovare una base di autovettori e scrivere la matrice che rappresenta la f in tale base.

Esercizio 7.5.18. Determinare autovalori e autovettori delle seguenti applicazioni lineari: $f : \mathbb{R}^2 \to \mathbb{R}^2$:

- $f(x, y) = (x + y, y)$;

- $f(x,y) = (x, x+y)$.

Esercizio 7.5.19. Determinare autovalori e autovettori delle seguenti applicazioni lineari: $f : \mathbb{R}^2 \to \mathbb{R}^2$:

- $f(x,y) = (2x - 3y, x)$,
- $f(x,y) = (x+y, x+y)$.

Esercizio 7.5.20. Determinare autovalori e autovettori della seguente applicazione lineare: $f : \mathbb{R}^2 \to \mathbb{R}^2$ con $f(x,y) = (x + ay, bx + y)$ dove a, b sono parametri reali.

Esercizio 7.5.21. Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare definita nel modo seguente: $f(x,y,z) = (y - z, -x + 2y - z, x - y + 2z)$. Dimostrare che è semplice e trovare una base di autovettori per $\mathbb{R}^3$ e determinare la matrice che rappresenta f in tale base.

Esercizio 7.5.22. Per ognuna delle applicazioni lineari $f : \mathbb{R}^3 \to \mathbb{R}^3$ date, stabilire se sia semplice. Descriverne poi gli autospazi.

- $f(x,y,z) = (-6x + 2y - 5z, -4x + 4y - 2z, 10x - 3y + 8z)$;
- $f(x,y,z) = (-8x - 13y - 14z, -6x - 5y - 8z, 14x + 17y + 21z)$;
- $f(x,y,z) = (-5x + 4y - z, -4x + y - 2z, 8x - 4y + 3z)$;

Esercizio 7.5.23. Per ognuna delle applicazioni lineari $f : \mathbb{R}^3 \to \mathbb{R}^3$ date, stabilire se sia semplice. Descriverne poi gli autospazi.

- $f(x,y,z) = (13x + 59y + 34z, 10x + 40y + 24z, -18x - 79y - 46z)$;
- $f(x,y,z) = (-4x + 2y + 5z, -44y - 120z, 16y + 44z)$;

Esercizio 7.5.24. Stabilire se l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ così definita

$$f(x,y,z) = (x, -3y - 15z, 2y + 8z)$$

è semplice. In caso affermativo trovare una base di autovettori per $\mathbb{R}^3$ e determinare la matrice che rappresenta f in tale base.

Esercizio 7.5.25. Stabilire se l'applicazione lineare $f : \mathbb{R}^4 \to \mathbb{R}^4$ così definita

$$f(x,y,z,t) = (y,z,t,x)$$

è semplice. Descriverne poi gli autospazi.

Esercizio 7.5.26. Trovare per quali valori reali di $a, b \in \mathbb{R}$ l'applicazione lineare $f : \mathbb{R}^2 \to \mathbb{R}^2$ così definita $f(x,y) = (x + ay, bx + y)$ è semplice.

Esercizio 7.5.27. Trovare per quali valori reali di $k \in \mathbb{R}$ l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ data è semplice.

- $f(x,y,z) = (kx + y, y + z, z)$;

- $f(x,y,z) = (y - kz, x - kz, kz)$;

Esercizio 7.5.28. Per i k che rendono semplici le applicazioni lineari dell'esercizio precedente disegnare autovettori, le loro rispettive immagini e gli autospazi.

Esercizio 7.5.29. Trovare per quali valori reali di $k \in \mathbb{R}$ l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ data è semplice.

- $f(x,y,z) = (2kz, x + ky, x)$;
- $f(x,y,z) = (y, x - ky, z)$;

Esercizio 7.5.30. Trovare per quali valori reali di $k \in \mathbb{R}$ l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ data è semplice.

- $f(x,y,z) = (x + y, y, kz)$;
- $f(x,y,z) = (10x, -ky + 3z, x + y - z)$.

Esercizio 7.5.31. Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$, definita da $f(6,9) = (2,3)$ e $f(4,2) = (2,1)$. La f è semplice?

Esercizio 7.5.32. Trovare se esistono dei valori di $k \in \mathbb{R}$ tali che $f : \mathbb{R}^3 \to \mathbb{R}^3$ con $f(x,y,z) = (x - ky, y - kz, z - x)$ sia semplice.

Esercizio 7.5.33. Scrivere una applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ tale che il sottospazio $S = \{(x,y,z) \in \mathbb{R}^3 \,|\, x - y = y + z = 0)\}$ ne sia un autospazio.

Esercizio 7.5.34. Scrivere una applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ tale che il sottospazio $S = \{(x,y,z) \in \mathbb{R}^3 \,|\, x + 3y - 4z = 0)\}$ ne sia un autospazio.

Esercizio 7.5.35. Quante applicazioni lineari $f : \mathbb{R}^3 \to \mathbb{R}^3$ è possibile scrivere in modo tale che ogni vettore $\mathbf{v} \in \mathbb{R}^3$ venga portato in $2\mathbf{v} \in \mathbb{R}^3$? Quanti autospazi possiede una tale applicazione? Di quali dimensioni?

Esercizio 7.5.36. * Trovare autovalori e autovettori per $A \in \mathbb{R}^{n,n}$, con:

$$A = hI_n + \begin{pmatrix} 0 & 0 & \dots & k \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ k & 0 & 0 & 0 \end{pmatrix}.$$

Dire, al variare di (h,k), se A sia diagonalizzabile.

Esercizio 7.5.37. Trovare autovalori e autovettori per

$$A = \begin{pmatrix} 0 & h & 0 \\ k & 0 & k \\ 0 & h & 0 \end{pmatrix}.$$

Dire, al variare di (h,k), se A sia diagonalizzabile.

Esercizio 7.5.38. * Sia $A \in \mathbb{R}^{2,2}$ una matrice simmetrica. Dimostrare che A ha solo autovalori reali e è sempre diagonalizzabile.

Esercizio 7.5.39. Sia $f_0 : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione nulla (ogni vettore viene portato da f_0 nel vettore nullo. Trovare autovalori e autovettori di f_0 e dire se f_0 sia semplice.

Esercizio 7.5.40. Disegnare immagine, nucleo e autospazi della seguente applicazione lineare:
$f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (-x - y, x + 2y - z, y - z)$.

CAPITOLO 8

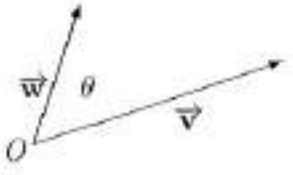
Spazi vettoriali euclidei

Prerequisiti: Spazi vettoriali, applicazioni lineari.

Obiettivi: Conoscenza della struttura di spazio vettoriale euclideo, saper operare con prodotti scalari anche non-standard; saper utilizzare basi ortonormali sfruttando le loro proprietà; determinare spazi ortogonali a sottospazi dati in uno spazio vettoriale euclideo.

8.1 Prodotto scalare e spazi vettoriali euclidei

Ricordiamo la definizione di prodotto scalare di due vettori del piano $\mathcal{V}_O^2$ (vale in modo del tutto analogo anche in $\mathcal{V}_O^3$).

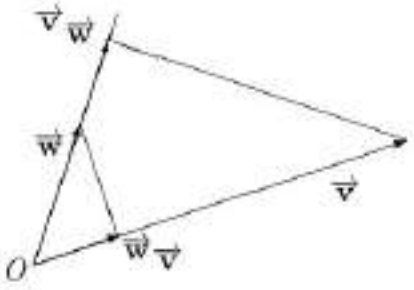


Definizione 8.1.1. Siano $\vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathcal{V}_O^2$ e sia θ l'angolo convesso tra di essi; allora si dice *prodotto scalare* tra $\vec{\mathbf{v}}$ e $\vec{\mathbf{w}}$ il numero:

$$\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = |\vec{\mathbf{v}}||\vec{\mathbf{w}}| \cos\theta.$$

Abbiamo quindi che l'operazione di prodotto scalare tra vettori è una funzione $\mathcal{V}_O^2 \times \mathcal{V}_O^2 \to \mathbb{R}$.

Geometricamente, se dati due vettori $\vec{\mathbf{z}}$, $\vec{\mathbf{u}} \in \mathcal{V}_O^2$, indichiamo con $\vec{\mathbf{z}}_{\vec{\mathbf{u}}}$ il vettore ottenuto dalla proiezione di $\vec{\mathbf{z}}$ nella direzione di $\vec{\mathbf{u}}$, allora si ha:



$$|\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}| = |\vec{\mathbf{v}}||\vec{\mathbf{w}}_{\vec{\mathbf{v}}}| = |\vec{\mathbf{v}}_{\vec{\mathbf{w}}}||\vec{\mathbf{w}}|;$$

Esempio 8.1.1. In fisica, il lavoro L compiuto da una forza costante $\vec{\mathbf{F}}$ lungo uno spostamento rettilineo rappresentato dal vettore $\vec{\mathbf{s}}$, è dato da $L = \vec{\mathbf{F}} \cdot \vec{\mathbf{s}}$ (quindi $|L| = |\vec{\mathbf{F}}_{\vec{\mathbf{s}}}||\vec{\mathbf{s}}|$: solo la componente $\vec{\mathbf{F}}_{\vec{\mathbf{s}}}$ della forza "compie il lavoro").

In particolare avremo:

$$\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = 0 \quad \text{se e solo se} \quad \vec{\mathbf{v}} \text{ e } \vec{\mathbf{w}} \text{ sono perpendicolari.}$$

Notiamo che il vettore nullo $\mathbf{0}$ si considera perpendicolare a ogni vettore di $\mathcal{V}_O^2$.

Come proprietà peculiari del prodotto scalare, si ha che:

- $\forall\, \vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathcal{V}_O^2,\ \vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = \vec{\mathbf{w}} \cdot \vec{\mathbf{v}}$;
- $\forall\, \vec{\mathbf{v}} \in \mathcal{V}_O^2,\quad \sqrt{\vec{\mathbf{v}} \cdot \vec{\mathbf{v}}} = |\vec{\mathbf{v}}| \geq 0$;
- Dal punto precedente si ha che $\vec{\mathbf{v}} \cdot \vec{\mathbf{v}} = 0$ se e solo se $\vec{\mathbf{v}} = \mathbf{0}$;
- $\forall\, \vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{w}} \in \mathcal{V}_O^2,\ (\vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2) \cdot \vec{\mathbf{w}} = \vec{\mathbf{v}}_1 \cdot \vec{\mathbf{w}} + \vec{\mathbf{v}}_2 \cdot \vec{\mathbf{w}}$;
- $\forall\, \vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathcal{V}_O^2, \forall \alpha \in \mathbb{R},\ (\alpha \vec{\mathbf{v}}) \cdot \vec{\mathbf{w}} = \vec{\mathbf{v}} \cdot (\alpha \vec{\mathbf{w}}) = \alpha(\vec{\mathbf{v}} \cdot \vec{\mathbf{w}})$.

Dato un sistema di assi cartesiani di origine O in $\mathcal{V}_O^2$, vediamo come si può rileggere il prodotto scalare quando si passa da $\mathcal{V}_O^2$ a $\mathbb{R}^2$ tramite l'isomorfismo che associa a ogni vettore $\overrightarrow{OP} \in \mathcal{V}_O^2$ le coordinate $(x_P, y_P) \in \mathbb{R}^2$.

Siano $\vec{\mathbf{i}}$ e $\vec{\mathbf{j}}$ i due versori (cioè vettori di modulo 1) corrispondenti ai punti $(1,0)$ e $(0,1)$ in $\mathbb{R}^2$ (quindi la base $\{\vec{\mathbf{i}}, \vec{\mathbf{j}}\}$ di $\mathcal{V}_O^2$ corrisponde

alla base canonica di $\mathbb{R}^2$); allora ogni vettore di $\mathcal{V}_O^2$ si può scrivere come combinazione lineare di $\vec{\imath}$ e $\vec{\jmath}$, cosicché se abbiamo due vettori:

$$\vec{\mathbf{v}} = a\,\vec{\imath} + b\,\vec{\jmath}, \qquad \vec{\mathbf{w}} = c\,\vec{\imath} + d\,\vec{\jmath},$$

il loro prodotto scalare sarà:

$$\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = (a\,\vec{\imath} + b\,\vec{\jmath}) \cdot (c\,\vec{\imath} + d\,\vec{\jmath}) = ac\,\vec{\imath} \cdot \vec{\imath} + ad\,\vec{\imath} \cdot \vec{\jmath} + bc\,\vec{\jmath} \cdot \vec{\imath} + bd\,\vec{\jmath} \cdot \vec{\jmath}$$

Poiché $\vec{\imath} \cdot \vec{\imath} = \vec{\jmath} \cdot \vec{\jmath} = 1$, mentre $\vec{\imath} \cdot \vec{\jmath} = \vec{\jmath} \cdot \vec{\imath} = 0$, ricaviamo che

$$\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = ac + bd.$$

Quindi se vogliamo lavorare in $\mathbb{R}^2$, si può rileggere da $\mathcal{V}_O^2$ il prodotto scalare come definito dalla seguente legge:

$$(a, b) \cdot (c, d) = ac + bd.$$

Analogamente, in $\mathbb{R}^3$, scegliendo una base $\{\vec{\imath}, \vec{\jmath}, \vec{\mathbf{k}}\}$ di $\mathcal{V}_O^3$, corrispondente a quella canonica di $\mathbb{R}^3$, avremmo ottenuto

$$(a_1, a_2, a_3) \cdot (b_1, b_2, b_3) = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Si vede quindi che si possono esprimere tramite il prodotto scalare varie proprietà geometriche degli "oggetti" di $\mathbb{R}^2$ (sempre via l'isomorfismo che identifica $\mathcal{V}_O^2$ e $\mathbb{R}^2$), come distanze, lunghezza di segmenti, perpendicolarità. Queste sono tutte nozioni che non hanno senso negli spazi vettoriali $\mathbb{R}^n$, in quanto sono nozioni che non sono esprimibili solo tramite le operazioni di somma e prodotto per gli scalari (le uniche presenti negli spazi vettoriali).

Quello che vogliamo fare è estendere la nozione di prodotto scalare a spazi vettoriali qualsiasi, in modo da usarlo per definire anche in altri ambienti delle nozioni di distanza, di angolo, di lunghezza.

Ricordando le principali proprietà che abbiamo notato per il prodotto scalare in $\mathcal{V}_O^2$, diamo la seguente definizione:

Definizione 8.1.2. Sia V uno spazio vettoriale su $\mathbb{R}$. Consideriamo una funzione $" \cdot ": V \times V \to \mathbb{R}$, che goda delle proprietà seguenti:

1. $\forall\, \mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in \mathbf{V}$: $\mathbf{v} \cdot (\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{v} \cdot \mathbf{v}_1 + \mathbf{v} \cdot \mathbf{v}_2$;
2. $\forall\, \mathbf{v}, \mathbf{w} \in \mathbf{V}$, $\forall \lambda \in \mathbb{R}$: $\mathbf{v} \cdot (\lambda \mathbf{w}) = (\lambda \mathbf{v}) \cdot \mathbf{w} = \lambda(\mathbf{v} \cdot \mathbf{w})$;
3. $\forall\, \mathbf{v}, \mathbf{w} \in V$: $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$;
4. $\forall\, \mathbf{v} \in V$: $\mathbf{v} \cdot \mathbf{v} \geq \mathbf{0}$, e $\mathbf{v} \cdot \mathbf{v} = \mathbf{0}$ se e solo se $\mathbf{v} = \mathbf{0}_V$.

Allora $" \cdot "$ si dice un *prodotto scalare* su V.

In questo modo, quelle che erano in $\mathcal{V}_O^2$ le proprietà del prodotto scalare ci hanno indicato come articolare la definizione stessa di "prodotto scalare" in spazi vettoriali qualsiasi, proprio come le proprietà di $\mathcal{V}_O^2$ sono servite da "prototipo" per la definizione della struttura di spazio vettoriale in generale.

Potremo quindi considerare adesso le strutture date da uno spazio vettoriale dotato di un prodotto scalare.

Definizione 8.1.3. Sia V uno spazio vettoriale su $\mathbb{R}$ dotato di un prodotto scalare. Allora la struttura $(V, \cdot)$ si dice uno *spazio vettoriale euclideo*.

Per esempio vediamo di definire un prodotto scalare in tutti gli spazi vettoriali $\mathbb{R}^n$; per farlo useremo come modello il modo in cui il prodotto scalare si legge in $\mathbb{R}^2$: $(a, b) \cdot (c, d) = ac + bd$.

Definizione 8.1.4. Siano $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, con $\mathbf{v} = (a_1, a_2, \ldots, a_n)$, mentre $\mathbf{w} = (b_1, b_2, \ldots, b_n)$; si pone:

$$\mathbf{v} \cdot \mathbf{w} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Tale operazione viene indicata come *prodotto scalare standard* in $\mathbb{R}^n$.

Pensando i vettori come matrici colonna e riga, si ha che il prodotto scalare standard coincide con il prodotto righe per colonne fra matrici:

$$(a_1\ a_2\ \ldots\ a_n) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Non è difficile provare che questo prodotto soddisfa i punti 1,...,4 della definizione precedente, e che quindi è in effetti un prodotto scalare (farlo per esercizio).

In particolare lo spazio vettoriale euclideo $(\mathbb{R}^n, \cdot)$, ove " $\cdot$ " è il prodotto scalare standard, si dice *spazio euclideo standard* (n-dimensionale) e si indica con $\mathbb{E}^n$.

Osserviamo che se $n = 2$ o $n = 3$, si ha che il prodotto scalare standard nello spazio $\mathbb{E}^n$ coincide con il prodotto scalare ordinario quando "riletto" in $\mathcal{V}_O^2$, $\mathcal{V}_O^3$.

Esempio 8.1.2. Su $\mathbb{R}^n$ si possono avere altri prodotti scalari oltre a quello standard, ad esempio, in $\mathbb{R}^4$ per il prodotto:

$$(a_1, a_2, a_3, a_4) \times (b_1, b_2, b_3, b_4) = a_1b_1 + 2a_2b_2 + 3a_3b_3 + 4a_4b_4;$$

è facile verificare che soddisfa le quattro condizioni per essere un prodotto scalare, ma $(\mathbb{R}^4, \times)$ è uno spazio vettoriale euclideo diverso da $\mathbb{E}^4$.

Esempio 8.1.3. Consideriamo lo spazio vettoriale $\mathcal{C}_{[a,b]}$ dato dalle funzioni continue $f : [a, b] \to \mathbb{R}$, definite su un intervallo $[a, b] \subset \mathbb{R}$. In $\mathcal{C}_{[a,b]}$ si può definire:

$$\forall\, f, g \in \mathcal{C}_{[a,b]} : \qquad f \cdot g = \int_a^b f(x)g(x)dx.$$

È facile verificare che " $\cdot$ " soddisfa le proprietà 1, 2 e 3 della Definizione 8.1.2. Per la 4, notiamo che banalmente si ha $f \cdot f \geq 0$, in quanto l'integrale della funzione f^2 sarà non negativo. Vedere che $f \cdot f = \mathbf{0}$ se e solo se $f = \mathbf{0}$ (la funzione nulla) è un po' più delicato. Notiamo che se esiste un punto $c \in [a, b]$ ove $f(c) \neq 0$, allora $f^2(c) > 0$, e per il teorema della permanenza del segno (vedi il corso di Analisi), c'è tutto un intorno $(c - \delta, c + \delta)$ ove la f^2 è positiva, quindi

$$f \cdot f \;=\; \int_a^b f^2(x)dx \geq \int_{c-\delta}^{c+\delta} f^2(x)dx > 0.$$

Quindi si può avere $f \cdot f = 0$ se e solo se $f = \mathbf{0}$. Si ha così che $\mathcal{C}_{[a,b]}$ è uno spazio vettoriale euclideo con il prodotto scalare " $\cdot$ ".

8.2 Norma, ortogonalità

In uno spazio vettoriale euclideo $(V, \cdot)$ si definisce la *norma* di un vettore:

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

La norma è l'analogo della nozione di *modulo* che abbiamo in $\mathcal{V}_O^2$, $\mathcal{V}_O^3$: la norma soddisfa le proprietà seguenti:

Proposizione 8.2.1. *Sia $(V, \cdot)$ uno spazio vettoriale euclideo. Allora $\forall\, \mathbf{v}, \mathbf{w} \in V$:*

- $||\mathbf{v}|| \geq 0$, *e* $||\mathbf{v}|| = 0$ *se e solo se* $\mathbf{v} = \mathbf{0}_V$;

- $\forall \alpha \in \mathbb{R}, \quad ||\alpha \mathbf{v}|| = |\alpha| ||\mathbf{v}||$;
- $\|\mathbf{v}+\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + 2\mathbf{v}\cdot\mathbf{w} + \|\mathbf{w}\|^2$;
- $\|\mathbf{v}\|\,\|\mathbf{w}\| \geq |\mathbf{v}\cdot\mathbf{w}|$ *(disuguaglianza di Cauchy-Schwartz)*;
- $|(\|\mathbf{v}\| - \|\mathbf{w}\|)| \leq \|\mathbf{v}+\mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ *(disuguaglianza triangolare)*;

Dimostrazione. • Il primo punto è conseguenza diretta delle proprietà del prodotto scalare (vedi Def. 8.1.2, punto 4).

- Per il secondo punto:
  $\|\alpha\mathbf{v}\| = \sqrt{(\alpha\mathbf{v})\cdot(\alpha\mathbf{v})} = \sqrt{\alpha^2(\mathbf{v}\cdot\mathbf{v})} = |\alpha|\sqrt{\mathbf{v}\cdot\mathbf{v}} = |\alpha|\|\mathbf{v}\|$.
- Terzo punto:
  $\|\mathbf{v}+\mathbf{w}\|^2 = (\mathbf{v}+\mathbf{w})\cdot(\mathbf{v}+\mathbf{w}) = \mathbf{v}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{w} + \mathbf{w}\cdot\mathbf{v} + \mathbf{w}\cdot\mathbf{w} = \|\mathbf{v}\|^2 + 2\mathbf{v}\cdot\mathbf{w} + \|\mathbf{w}\|^2$. Notiamo che se $\mathbf{v}$ e $\mathbf{w}$ sono ortogonali, questo punto è equivalente al Teorema di Pitagora in $\mathcal{V}_O^2$.
- Disuguaglianza di Cauchy–Schwartz:
  Se uno dei due vettori è nullo, vale banalmente l'uguaglianza. Quindi supponiamo che entrambi siano non nulli e consideriamo, $\forall \alpha, \beta \in \mathbb{R}$:
  $$0 \leq \|\alpha\mathbf{v} + \beta\mathbf{w}\|^2 = \alpha^2\|\mathbf{v}\| + 2\alpha\beta\mathbf{v}\cdot\mathbf{w} + \beta^2\|\mathbf{w}\|^2;$$
  se poniamo $\alpha = \|\mathbf{w}\|$ e $\beta = -\mathbf{v}\cdot\mathbf{w}$, otteniamo:
  $$0 \leq \|\mathbf{w}\|^2(\|\mathbf{v}\|^2\|\mathbf{w}\|^2 - (\mathbf{v}\cdot\mathbf{w})^2)$$
  da cui, dividendo per $\|\mathbf{w}\|^2$, si ricava la disuguaglianza voluta.
- Disuguaglianza triangolare:
  dal punto precedente si ha:
  $$(\|\mathbf{v}\| - \|\mathbf{w}\|)^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\| \leq \|\mathbf{v}\|^2 + 2\mathbf{v}\cdot\mathbf{w} + \|\mathbf{w}\|^2 =$$
  $$= \|\mathbf{v}+\mathbf{w}\|^2 \leq \|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2 = (\|\mathbf{v}\| + \|\mathbf{w}\|)^2$$
  dalla quale segue la diseguaglianza triangolare.

□

La precedente proposizione mostra come la norma si comporti come la lunghezza dei vettori nella geometria euclidea classica: ogni vettore ha lunghezza positiva o nulla, e solo il vettore nullo ha lunghezza nulla; il multiplo di un vettore $\vec{\mathbf{v}}$ ha lunghezza pari a quella di $\vec{\mathbf{v}}$ moltiplicata per il valore assoluto del numero per cui lo moltiplichiamo; la disuguaglianza

triangolare è l'analogo della proprietà dei triangoli per cui ogni lato è maggiore della differenza degli altri due ed è minore della loro somma.

La seguente definizione mostra come si possono estendere a ogni spazio vettoriale euclideo gli usuali concetti di *distanza*, *perpendicolarità* e *angolo*.

Definizione 8.2.1. Sia $(V, \cdot)$ uno spazio vettoriale euclideo.

- $\forall\, \mathbf{v}, \mathbf{w} \in V$, si definisce il numero $d(\mathbf{v}, \mathbf{w})$ come *distanza* fra $\mathbf{v}$ e $\mathbf{w}$, ponendo $d(\mathbf{v}, \mathbf{w}) = ||\mathbf{v} - \mathbf{w}||$;
- Si definisce l'*angolo* θ fra i due vettori con:

$$\theta = \arccos \frac{\mathbf{v} \cdot \mathbf{w}}{||\mathbf{v}||\ ||\mathbf{w}||},$$

  avremo allora: $\mathbf{v} \cdot \mathbf{w} = ||\mathbf{v}||\ ||\mathbf{w}|| \cos\theta$;
- In particolare, $\mathbf{v}, \mathbf{w} \in V$ si dicono *ortogonali* o *perpendicolari*, e si scrive $\mathbf{v} \perp \mathbf{w}$, se e solo se: $\mathbf{v} \cdot \mathbf{w} = 0$.

Notiamo alcune cose:

La definizione dell'angolo θ è ben posta, in quanto la frazione utilizzata è sempre ≤ 1 in valore assoluto (per la diseguaglianza di Cauchy–Schwartz).

Il vettore nullo risulta l'unico che è ortogonale ad ogni altro vettore (e anche a se stesso). Infatti se un vettore $\mathbf{v} \in V$ è ortogonale a ogni altro, si avrebbe, preso $\mathbf{w} \neq \mathbf{v}$:

$$\mathbf{v} \cdot \mathbf{v} = \mathbf{v} \cdot [(\mathbf{v} - \mathbf{w}) + \mathbf{w}] = \mathbf{v} \cdot (\mathbf{v} - \mathbf{w}) + \mathbf{v} \cdot \mathbf{w} = 0,$$

da cui segue $\mathbf{v} = \mathbf{0}_V$.

Vettori ortogonali fra loro sono anche linearmente indipendenti:

Proposizione 8.2.2. *Sia $(V, \cdot)$ uno spazio vettoriale euclideo e sia $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subset V$ un insieme di vettori ortogonali (cioè vettori tali che $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, $\forall\, i \neq j$) e non nulli. Allora i $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono linearmente indipendenti.*

Dimostrazione. Consideriamo una combinazione lineare che dia il vettore nullo: $a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}_V$; vogliamo vedere che necessariamente i coefficienti a_i devono essere tutti nulli.

Poiché i vettori $\mathbf{v}_i$ non sono nulli, $\forall\, i = 1, \ldots, n$ avremo:

$$\mathbf{v}_i \cdot (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = \mathbf{v}_i \cdot \mathbf{0}_V = 0,$$

ma $\mathbf{v}_i \cdot (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = a_i||\mathbf{v}_i||^2$, quindi, per ogni $i = 1, \ldots, n$, vale $a_i = 0$, e abbiamo concluso. □

Tutto ciò spiega perchè chiamiamo "euclidei" questi spazi: grazie alla presenza di un prodotto scalare è possibile definirvi concetti analoghi a quelli della geometria euclidea (angoli, perpendicolarità, lunghezze, distanze), e utilizzarli anche in questa accezione così generalizzata.

Esempio 8.2.1. Siano $\mathbf{v} = (a_1, \ldots, a_n)$ e $\mathbf{w} = (b_1, \ldots, b_n)$ in $\mathbb{E}^n$; allora

$$d(\mathbf{v}, \mathbf{w}) = ||\mathbf{v} - \mathbf{w}|| = \sqrt{(a_1 - b_1, \ldots, a_n - b_n) \cdot (a_1 - b_1, \ldots, a_n - b_n)} =$$

$$= \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2},$$

che estende a tutti gli $\mathbb{E}^n$ la ben nota formula per la distanza fra due punti incontrata nello studio della geometria analitica del piano.

Esempio 8.2.2. Siano $\mathbf{v} = (2, 0, 1, k)$ e $\mathbf{w} = (0, 1, k, -1)$ in $\mathbb{E}^4$; determinare la distanza $d(\mathbf{v}, \mathbf{w})$ in funzione di k e dire se esiste un valore di k per cui $\mathbf{v} \perp \mathbf{w}$. Si ha (dall'esempio precedente):

$$d(\mathbf{v}, \mathbf{w}) = \sqrt{(2-0)^2 + (0-1)^2 + (1-k)^2 + (k+1)^2} =$$

$$= \sqrt{4 + 1 + 1 - 2k + k^2 + k^2 + 2k + 1} = \sqrt{2k^2 + 7}.$$

Per quanto riguarda la perpendicolarità, basta porre $\mathbf{v} \cdot \mathbf{w} = 0$, che diventa:

$$\mathbf{v} \cdot \mathbf{w} = 2.0 + 0.1 + 1.k + k(-1) = 0,$$

quindi i due vettori risultano perpendicolari per ogni valore di k.

Definizione 8.2.2. Se $(V, \cdot)$ è uno spazio vettoriale euclideo, e $\mathbf{v} \in V$, si definisce *l'ortogonale* di $\mathbf{v}$ come:

$$\mathbf{v}^\perp = \{\mathbf{w} \in V |\ \mathbf{v} \cdot \mathbf{w} = 0\}.$$

Se poi $W \subset V$ è un sottospazio di V, diremo che $\mathbf{v} \in V$ è *ortogonale* a W (scrivendo $\mathbf{v} \perp W$) quando $\forall \mathbf{w} \in \mathbf{W}$, $\mathbf{v} \cdot \mathbf{w} = 0$. Si definisce poi l'*ortogonale di* W:

$$W^\perp = \{\mathbf{v} \in V |\ \mathbf{v} \perp W\}.$$

Proposizione 8.2.3. *Sia $(V, \cdot)$ uno spazio vettoriale euclideo, $\forall\, \mathbf{v} \in V$ e per ogni sottospazio $W \subset V$, si ha che $\mathbf{v}^\perp$ e $W^\perp$ sono due sottospazi di V.*

Dimostrazione. Vediamolo per $W^\perp$. Si ha ovviamente $\mathbf{0}_V \in W^\perp$.

Se poi $\mathbf{v}_1, \mathbf{v}_2 \in W^\perp$ e $\alpha, \beta \in \mathbb{R}$, allora

$$\forall\, \mathbf{w} \in W, \quad (\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) \cdot \mathbf{w} = \alpha\mathbf{v}_1 \cdot \mathbf{w} + \beta\mathbf{v}_2 \cdot \mathbf{w} = 0$$

quindi $(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) \in W^\perp$ e $W^\perp$ è un sottospazio.

La dimostrazione per $\mathbf{v}^\perp$ è analoga ed è lasciata per esercizio. □

Osservazione 8.2.1. Notiamo che se $\mathcal{B} = \{\mathbf{w}_1, \ldots, \mathbf{w}_s\}$ è una base per il sottospazio W, allora

$$W^\perp = \{\mathbf{v} \in V \mid \mathbf{v} \cdot \mathbf{w}_i = 0,\ \forall i = 1, \ldots, s\}.$$

Infatti $\forall\, \mathbf{w} \in W$, $\mathbf{w} = a_1\mathbf{w}_1 + \cdots + a_s\mathbf{w}_s$, quindi se $\mathbf{v}$ è ortogonale a ogni $\mathbf{w}_i$, si ha anche $\mathbf{v} \cdot \mathbf{w} = a_1\mathbf{v} \cdot \mathbf{w}_1 + \cdots + a_s\mathbf{v} \cdot \mathbf{w}_s = 0$.

Esempio 8.2.3. Sia

$$W = \{(x, y, z) \in \mathbb{E}^3 \mid x - y + z = 0,\ x - z = 0\} \subset \mathbb{E}^3.$$

Qual è la dimensione di $W^\perp$? È facile vedere che $W = \{(z, 2z, z) \mid z \in \mathbb{R}\}$, quindi $\dim W = 1$, e una base di W è data dal vettore $\mathbf{w} = (1, 2, 1)$. Quindi i vettori di $W^\perp$ sono tutti e soli quelli perpendicolari a $\mathbf{w}$, cioè gli (x, y, z) tali che $(x, y, z) \cdot (1, 2, 1) = x + 2y + z = 0$.

Quindi $W^\perp = \{(-2y - z, y, z)\}$, $\dim W^\perp = 2$ e una sua base è data da $\mathcal{B} = \{(-2, 1, 0), (-1, 0, 1)\}$.

Geometricamente si ha che $W^\perp$ è un piano (mentre W è una retta).

Osservazione 8.2.2. Supponiamo che $W \subset \mathbb{E}^n$ sia un sottospazio definito da un sistema lineare omogeneo $A \cdot X = O$, ove $A \in \mathbb{R}^{m,n}$ e $X = {}^t(x_1\ x_2\ \ldots\ x_n)$. Osserviamo che i vettori di W, essendo le soluzioni del sistema, sono quelli ortogonali alle righe di A; quindi le righe di A sono generatori di $W^\perp$ e ovviamente $\dim W^\perp = r(A)$. Una base di $W^\perp$ è data da r righe indipendenti di A. Inoltre le soluzioni del sistema sono le stesse di quelle del sistema $A' \cdot X = O$, ove A' è ottenuta tramite eliminazione di Gauss su A, quindi, se si preferisce, si possono considerare le righe non nulle di A' per ottenere una base di $W^\perp$.

Esempio 8.2.4. Riprendendo l'esempio precedente, si ha che W è definito, in $\mathbb{E}^3$, dal sistema

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Quindi una base di $W^\perp$ è data da $\{(1, -1, 1), (1, 0, -1)\}$ (e $\dim W^\perp = 2$).

8.3 Basi ortonormali

In questa sezione esamineremo le basi che in uno spazio vettoriale euclideo è più conveniente usare, e cioè le basi ortonormali:

Definizione 8.3.1. Sia $(V, \cdot)$ uno spazio vettoriale euclideo, e $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ suoi vettori; essi si dicono *ortonormali* se

$$\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}.$$

Se $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base per V, e i vettori sono ortonormali, allora $\mathcal{B}$ si dice una *base ortonormale*.

Richiediamo quindi che i vettori della base siano ortogonali (cioè $\mathbf{v}_i \cdot \mathbf{v}_j = 0$) e abbiano tutti norma uno ($\mathbf{v}_i \cdot \mathbf{v}_i = 1$); i vettori di norma uno si dicono anche **versori**.

Se per una base $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ vale solo che i vettori sono ortogonali (senza essere necessariamente anche dei versori), si dice anche che è una *base ortogonale*.

Esempio 8.3.1. È immediato verificare che in $\mathbb{E}^n$ la base canonica $\mathcal{E}_n$ è una base ortonormale.

Esempio 8.3.2. In $\mathbb{E}^2$ la base $\mathcal{B} = \left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \right\}$ è una base ortonormale (verificare per esercizio).

In uno spazio vettoriale euclideo le basi ortonormali sono le più semplici da usare, infatti se esprimiamo un qualsiasi vettore $\mathbf{v} \in V$ rispetto a una base ortonormale $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$:

$$\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_n \mathbf{v}_n,$$

allora $\forall\, j = 1, \dots, n$ avremo

$$\mathbf{v} \cdot \mathbf{v}_j = (a_1 \mathbf{v}_1 + \cdots + a_n \mathbf{v}_n) \cdot \mathbf{v}_j = a_1 \mathbf{v}_1 \cdot \mathbf{v}_j + \cdots + a_j \mathbf{v}_j \cdot \mathbf{v}_j + \cdots + a_n \mathbf{v}_n \cdot \mathbf{v}_j = a_j,$$

cioè le coordinate a_j di $\mathbf{v}$ rispetto ai $\mathbf{v}_j$ sono esattamente i prodotti scalari $\mathbf{v} \cdot \mathbf{v}_j$.

Come conseguenza dell'osservazione precedente, abbiamo:

Proposizione 8.3.1. *Sia $(V, \cdot)$ uno spazio vettoriale euclideo e consideriamo $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, una sua base ortonormale. Allora $\forall\, \mathbf{v}, \mathbf{w} \in V$, se $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, $\mathbf{w} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$, avremo*

$$\mathbf{v} \cdot \mathbf{w} = a_1 b_1 + \cdots + a_n b_n.$$

La proposizione mostra che dato un qualsiasi spazio vettoriale euclideo V di dimensione n, il fissare una base ortogonale ci permette di lavorare in V come se fossimo in $\mathbb{E}^n$ (il prodotto scalare si esprime come quello standard).

Il seguente teorema garantisce l'esistenza di basi ortonormali e la possibilità di completare insiemi di vettori ortonormali a una base ortonormale.

Teorema 8.3.1. *Sia $(V, \cdot)$ uno spazio vettoriale euclideo non nullo, allora:*

- *$(V, \cdot)$ possiede una base ortonormale.*
- *ogni insieme di vettori ortonormali $\mathbf{w}_1, \ldots, \mathbf{w}_s$ in V si può completare a una base ortonormale di V.*

La dimostrazione di questo teorema nel caso di dimensione finita, che è l'unico che considereremo qui, è conseguenza immediata della seguente proposizione.

Proposizione 8.3.2. (Ortogonalizzazione di Gram–Schmidt) *Sia $(V, \cdot)$ uno spazio vettoriale euclideo e siano $\mathbf{v}_1, \ldots, \mathbf{v}_s \in V$ dei vettori linearmente indipendenti; allora esistono $\mathbf{w}_1, \ldots, \mathbf{w}_s \in V$ tali che:*

- *$\mathbf{w}_1, \ldots, \mathbf{w}_s$ sono ortonormali;*
- *$\forall\, k = 1, \ldots, s$, $\langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle = \langle \mathbf{w}_1, \ldots, \mathbf{w}_k \rangle$.*

Dimostrazione. I vettori $\mathbf{w}_i$ possono esser costruiti come segue:

$$\mathbf{w}_1 = \frac{\mathbf{v}_1}{||\mathbf{v}_1||}.$$

È immediato notare che $\langle \mathbf{w}_1 \rangle = \langle \mathbf{v}_1 \rangle$ e che

$$||\mathbf{w}_1|| = \left|\left| \frac{\mathbf{v}_1}{||\mathbf{v}_1||} \right|\right| = \frac{||\mathbf{v}_1||}{||\mathbf{v}_1||} = 1.$$

Poniamo poi

$$\widehat{\mathbf{w}}_2 = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{w}_1)\mathbf{w}_1;$$

e definiamo

$$\mathbf{w}_2 = \frac{\widehat{\mathbf{w}}_2}{||\widehat{\mathbf{w}}_2||}.$$

Si prosegue così definendo, $\forall\, k = 2, \ldots, n$:

$$\widehat{\mathbf{w}}_k = \mathbf{v}_k - (\mathbf{v}_k \cdot \mathbf{w}_1)\mathbf{w}_1 - (\mathbf{v}_k \cdot \mathbf{w}_2)\mathbf{w}_2 - \cdots - (\mathbf{v}_k \cdot \mathbf{w}_{k-1})\mathbf{w}_{k-1};$$

e poi

$$\mathbf{w}_k = \frac{\widehat{\mathbf{w}}_k}{||\widehat{\mathbf{w}}_k||}.$$

Vediamo che $\forall\, k = 1, \ldots, s$, i $\mathbf{w}_1, \ldots, \mathbf{w}_k$ sono ortonormali. Ogni $\mathbf{w}_i$ ha norma 1, in quanto

$$||\mathbf{w}_i|| = \left|\left| \frac{\widehat{\mathbf{w}}_i}{||\widehat{\mathbf{w}}_i||} \right|\right| = \frac{||\widehat{\mathbf{w}}_i||}{||\widehat{\mathbf{w}}_i||} = 1.$$

Rimane quindi da vedere che i $\mathbf{w}_i$ sono ortogonali fra loro. Procediamo per induzione su k.

- Se $k = 2$, abbiamo che $\{\mathbf{w}_1, \mathbf{w}_2\}$ è un insieme ortogonale, in quanto

$$\mathbf{w}_2 \cdot \mathbf{w}_1 = \frac{[\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{w}_1)\mathbf{w}_1]}{||\widehat{\mathbf{w}}_2||} \cdot \mathbf{w}_1 = \frac{1}{||\widehat{\mathbf{w}}_2||}[(\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{w}_1)\mathbf{w}_1) \cdot \mathbf{w}_1] =$$

$$= \frac{1}{||\widehat{\mathbf{w}}_2||}[\mathbf{v}_2 \cdot \mathbf{w}_1 - (\mathbf{v}_2 \cdot \mathbf{w}_1)(\mathbf{w}_1 \cdot \mathbf{w}_1)] = 0.$$

- Vediamo ora che se $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ è un insieme ortogonale, anche $\{\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{w}_{k+1}\}$ lo è. Dobbiamo verificare che $\mathbf{w}_{k+1} \cdot \mathbf{w}_i = 0$, $\forall\, i = 1, \ldots, k$. Avremo:

$$\mathbf{w}_{k+1} \cdot \mathbf{w}_i = \frac{1}{||\widehat{\mathbf{w}}_{k+1}||}(\widehat{\mathbf{w}}_{k+1} \cdot \mathbf{w}_i) =$$

$$= \frac{[\mathbf{v}_{k+1} - (\mathbf{v}_{k+1} \cdot \mathbf{w}_1)\mathbf{w}_1 - \cdots - (\mathbf{v}_{k+1} \cdot \mathbf{w}_k)\mathbf{w}_k] \cdot \mathbf{w}_i}{||\widehat{\mathbf{w}}_{k+1}||} =$$

Poiché abbiamo supposto che i $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ siano ortogonali, nell'espressione precedente l'unico prodotto scalare non nullo darà

$$= \frac{[\mathbf{v}_{k+1} \cdot \mathbf{w}_i - (\mathbf{v}_{k+1} \cdot \mathbf{w}_i)(\mathbf{w}_i \cdot \mathbf{w}_i)]}{||\widehat{\mathbf{w}}_{k+1}||} = \frac{\mathbf{v}_{k+1} \cdot \mathbf{w}_i - \mathbf{v}_{k+1} \cdot \mathbf{w}_i}{||\widehat{\mathbf{w}}_{k+1}||} = 0.$$

Quindi i due passi del procedimento di induzione sono verificati e ogni $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ è un insieme ortogonale.

Si ha poi che $\forall\, k = 1, \ldots s$, $\langle \mathbf{w}_1, \ldots, \mathbf{w}_k \rangle \subset \langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle$, in quanto ogni $\mathbf{w}_i$ è costruito come combinazione lineare dei $\mathbf{v}_1, \ldots, \mathbf{v}_i$. Inoltre

poiché sia i vettori $\mathbf{w}_1, \ldots, \mathbf{w}_k$ che $\mathbf{v}_1, \ldots, \mathbf{v}_k$ sono linearmente indipendenti, avremo che $\dim(\mathbf{w}_1, \ldots, \mathbf{w}_k) = \dim(\mathbf{v}_1, \ldots, \mathbf{v}_k) = k$, e quindi

$$\langle \mathbf{w}_1, \ldots \mathbf{w}_k \rangle = \langle \mathbf{v}_1, \ldots \mathbf{v}_k \rangle.$$

□

Ora si può dimostrare il Teorema 8.3.1, in quanto per Proposizione 2.3.3, lo spazio V ammette una base, e per la Proposizione 8.3.2 da quella base se ne può ricavare una ortonormale. Inoltre se ho un insieme di vettori ortonormali $\{\mathbf{w}_1, \ldots, \mathbf{w}_s\}$ in V, per la Proposizione 2.3.5 posso completarlo a una base $\{\mathbf{w}_1, \ldots, \mathbf{w}_s, \mathbf{v}_{s+1}, \ldots, \mathbf{v}_n\}$; si può poi da questo insieme ricavare una base ortonormale di V con il procedimento di Gram-Schmidt utilizzato nella Proposizione 8.3.2.

Come conseguenza di quanto visto finora si ha la seguente

Proposizione 8.3.3. *Sia $(V, \cdot)$ uno spazio vettoriale euclideo, di dimensione finita $\dim V = n$ e sia W un suo sottospazio. Allora si ha:*

1. $\dim W^\perp = \dim V - \dim W$;
2. $(W^\perp)^\perp = W$.

Dimostrazione. Per la prima parte: sia $\dim W = r$, e sia $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ una base ortonormale di W. Siano $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$, $n - r$ vettori che costituiscono un suo completamento a una base ortonormale di V (che esistono per il Teorema 8.3.1). Allora si ha che $\mathbf{v}_i \perp W, \forall\, i = r+1, \ldots, n$; quindi lo spazio $\langle \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n \rangle$ è incluso in $W^\perp$. D'altro canto, sia $\mathbf{v} \in V$, con

$$\mathbf{v} = a_1\mathbf{w}_1 + \cdots + a_r\mathbf{w}_r + a_{r+1}\mathbf{v}_{r+1} + \cdots + a_n\mathbf{v}_n,$$

se $\mathbf{v} \in W^\perp$, si avrà , per $i = 1, \ldots, r$, che $\mathbf{v} \cdot \mathbf{w}_i = a_i$, ma anche che $\mathbf{v} \cdot \mathbf{w}_i = 0$, e quindi $\mathbf{v} = a_{r+1}\mathbf{v}_{r+1} + \cdots + a_n\mathbf{v}_n$, cioè $\mathbf{v} \in \langle \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n \rangle$.

Quindi $W^\perp \subset \langle \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n \rangle$ e, avendo già notato l'inclusione opposta, si ha: $\langle \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n \rangle = W^\perp$. Poiché i $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ sono indipendenti, formano una base di $W^\perp$, cosicché $\dim W^\perp = n - r$.

Per dimostrare la seconda parte, notiamo che $W \subset (W^\perp)^\perp$ (dimostrarlo per esercizio). Inoltre, dalla prima parte, si ha che

$$\dim(W^\perp)^\perp = n - \dim W^\perp = n - (n - \dim W) = \dim W.$$

Quindi $W \subset (W^\perp)^\perp$ e hanno la stessa dimensione, perciò $W = (W^\perp)^\perp$.

□

8.4 Isometrie

Sia ora $f : V \to W$ un'applicazione lineare fra due spazi vettoriali euclidei V e W; poiché la f conserva la somma e la moltiplicazione per gli scalari, chiediamoci se la f conservi anche il prodotto scalare, e cioè se valga la seguente uguaglianza:

$$\mathbf{v} \cdot \mathbf{w} = f(\mathbf{v}) \cdot f(\mathbf{w}).$$

in generale la risposta è "NO".

Esempio 8.4.1. Sia $f : \mathbb{E}^2 \to \mathbb{E}^2$, con $f(x, y) = (3x, 2y)$; la f è ovviamente lineare, ma, ad esempio:

$$(2, 6) \cdot (1, 1) = 8, \text{ mentre } f(2, 6) \cdot f(1, 1) = (6, 12) \cdot (3, 2) = 42.$$

Non possiamo quindi dire che tutte le applicazioni lineari conservino il prodotto scalare. Vediamo allora se ne esistono alcune che lo conservano e cerchiamo di studiare quali sono le loro particolarità.

Osserviamo innanzitutto che se $f : V \to W$ conserva il prodotto scalare, allora deve essere iniettiva, infatti se $\ker f \neq \{0_V\}$, allora deve esistere $\mathbf{v} \in \ker f$, tale che $\mathbf{v} \neq 0_V$ e $f(\mathbf{v}) = 0_W$; ma allora avremmo $\mathbf{v} \cdot \mathbf{v} > 0$, mentre $f(\mathbf{v}) \cdot f(\mathbf{v}) = 0_W \cdot 0_W = 0$.

Ci concentreremo adesso sul caso in cui $f : V \to V$ sia un endomorfismo, e $\dim V = n$ (dimensione finita). Abbiamo già visto che per conservare il prodotto scalare f deve essere iniettiva, e quindi (per la Proposizione 5.3.5) è un isomorfismo. Non tutti gli isomorfismi conservano però il prodotto scalare (ad esempio la f dell'esempio precedente è un isomorfismo).

Definizione 8.4.1. Sia $(V, \cdot)$ uno spazio vettoriale euclideo e $f : V \to V$ un endomorfismo di V. Si dice che f è un'*isometria* se $\forall \mathbf{v}, \mathbf{w} \in V$ si ha $\mathbf{v} \cdot \mathbf{w} = f(\mathbf{v}) \cdot f(\mathbf{w})$ (cioè se f conserva il prodotto scalare).

Vediamo adesso, come si comportano le isometrie rispetto alle basi ortonormali.

Proposizione 8.4.1. *Un endomorfismo $f : V \to V$, ove $(V, \cdot)$ è uno spazio euclideo, è un'isometria se e solo se porta basi ortonormali in basi ortonormali.*

Dimostrazione. Come altre volte, eseguiremo la dimostrazione solo nel caso in cui $\dim V = n$ (dimensione finita).

Se f è un'isometria, allora per ogni base ortonormale $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, avremo

$$f(\mathbf{v}_i) \cdot f(\mathbf{v}_j) = \delta_{ij},$$

quindi i vettori $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)$ sono ortonormali e quindi linearmente indipendenti; essendo $n = \dim V$ vettori indipendenti, possono essere usati per formare una base. Quindi f porta basi ortonormali in basi ortonormali.

Supponiamo ora che f porti basi ortonormali in basi ortonormali e consideriamo una base ortonormale $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Comunque presi due vettori $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ e $\mathbf{w} = b_1\mathbf{v}_1 + \ldots + b_n\mathbf{v}_n \in V$, avremo (dalla Proposizione 8.3.1):

$$\mathbf{v} \cdot \mathbf{w} = a_1b_1 + \cdots + a_nb_n.$$

Considerando poi $f(\mathbf{v}) = a_1f(\mathbf{v}_1) + \cdots + a_nf(\mathbf{v}_n)$, $f(\mathbf{w}) = b_1f(\mathbf{v}_1) + \cdots + b_nf(\mathbf{v}_n)$, avremo, ancora per la Proposizione 8.3.1 poiché i vettori $\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)\}$ formano una base ortonormale:

$$f(\mathbf{v}) \cdot f(\mathbf{w}) = a_1b_1 + \cdots + a_nb_n.$$

Quindi $\mathbf{v} \cdot \mathbf{w} = f(\mathbf{v}) \cdot f(\mathbf{w})$, e la f è un'isometria. □

Allora, se f è un'isometria, sia $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base ortonormale, e $A = M_{\mathcal{B},\mathcal{B}}(f)$; le colonne di A sono i vettori $f(\mathbf{v}_i)$ (espressi rispetto alla base $\mathcal{B}$), e la base $\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)\}$ risulterà ortonormale.

Consideriamo allora il prodotto:

$${}^tA \cdot A = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = B.$$

L'elemento b_{ij} della matrice prodotto è il prodotto della riga i-esima di ${}^tA$ (che altro non è che la i-esima colonna di A) per la j-sima colonna di A; esse rappresentano il vettore $f(\mathbf{v}_i)$ e il vettore $f(\mathbf{v}_j)$ (rispetto alla base $\mathcal{B}$) e quindi si avrà (per la Proposizione 8.3.1):

$$b_{ij} = (a_{1i}, a_{2i}, \ldots, a_{ni}) \cdot \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = f(\mathbf{v}_i) \cdot f(\mathbf{v}_j) = \delta_{ij}$$

e quindi ${}^tA \cdot A = I$, perché avrà degli 1 sulla diagonale (i b_{ii}=1) e degli 0 altrove (i $b_{ij} = 0$ per $i \neq j$).

Definizione 8.4.2. Le matrici $A \in \mathbb{R}^{n,n}$ per cui si abbia ${}^tA \cdot A = I$ (cioè ${}^tA = A^{-1}$) sono dette *matrici ortogonali*.

Con le considerazioni precedenti abbiamo allora dimostrato la seguente:

Proposizione 8.4.2. *Sia $f : V \to V$, un endomorfismo, ove V è uno spazio euclideo e $\dim V = n$. La condizione necessaria e sufficiente perché f sia un'isometria è che comunque scelta una base ortonormale $\mathcal{B}$, la matrice $A = M_{\mathcal{B}\mathcal{B}}(f)$, che rappresenta la f rispetto a $\mathcal{B}$, sia* **ortogonale** *(cioè valga ${}^tA \cdot A = I$).*

Esempio 8.4.2. Consideriamo qui il caso delle isometrie $f : \mathbb{E}^2 \to \mathbb{E}^2$, cioè gli isomorfismi del piano euclideo che conservano le distanze e gli angoli. Per quanto visto, la matrice $A = M_{\mathcal{E},\mathcal{E}}(f)$, che rappresenta la f rispetto alla base canonica $\mathcal{E}$ (che è ortonormale), sarà una matrice ortogonale, quindi si ha:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

da cui:

$$\begin{cases} a^2+b^2=1 \\ ac+bd=0 \\ c^2+d^2=1 \end{cases}.$$

Ad esempio, consideriamo il caso della f_x definita dalla matrice con $a = -d = 1$, $b = c = 0$. Avremo:

$$f_x(x,y) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix};$$

quindi, geometricamente, la f_x è una *riflessione* che ha come asse di simmetria l'asse delle x.

In generale, dalla prima e terza equazione si ha che esistono θ, ϕ tali che: $a = \cos\theta$, $b = \sin\theta$ e $d = \cos\varphi, c = \sin\varphi$, cosicché la seconda equazione ci dà $\cos\theta\sin\varphi + \sin\theta\cos\varphi = \sin(\theta+\varphi) = 0$, per cui $\varphi = -\theta$ oppure $\varphi = \pi - \theta$. Le due possibili scelte ci portano alla conclusione che le isometrie cercate sono di due tipi: f_θ, e g_θ, ove

$$M_{\mathcal{E},\mathcal{E}}(f_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad M_{\mathcal{E},\mathcal{E}}(g_\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix},$$

Interpretiamo geometricamente i due casi;

$$f_\theta(x,y) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta x - \sin\theta y \\ \sin\theta x + \cos\theta y \end{pmatrix}.$$

Se poniamo $(x,y) = (\rho\cos\alpha, \rho\sin\alpha)$ usando le coordinate polari, vedremo $P = (x,y)$ come il punto che dista ρ dall'origine e tale che la semiretta OP forma un angolo α con l'asse delle x. Avremo allora che

$$f_\theta(x,y) = (\rho\cos\theta\cos\alpha - \rho\sin\theta\sin\alpha,\ \rho\sin\theta\cos\alpha + \rho\cos\theta\sin\alpha) =$$

$$= (\rho\cos(\alpha+\theta),\ \rho\sin(\alpha+\theta));$$

cioè la f_θ porta ogni punto $(x,y) = (\rho\cos\alpha, \rho\sin\alpha)$ nel punto $f_\theta(x,y) = (\rho\cos(\alpha+\theta),\ \rho\sin(\alpha+\theta))$ che si ottiene tramite la rotazione di un angolo θ intorno all'origine.

Per interpretare le g_θ, osserviamo che nel caso $\theta = 0$, la matrice associata a g_0 è quella vista all'inizio, cioè si ha $g_0 = f_x$: la g_0 è la riflessione del piano rispetto all'asse x. Notiamo poi che:

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

e cioè che

$$M_{\mathcal{E},\mathcal{E}}(g_\theta) = M_{\mathcal{E},\mathcal{E}}(f_\theta) \cdot M_{\mathcal{E},\mathcal{E}}(g_0).$$

Quindi $g_\theta = f_\theta \circ f_x$, cosicché l'effetto di g_θ è quello di prendere un punto (x,y), farne il simmetrico rispetto all'asse x e poi ruotarlo di un angolo θ intorno all'origine.

8.5 Il teorema spettrale

Abbiamo visto nel capitolo precedente che una matrice quadrata $A \in \mathbb{R}^{n,n}$ è diagonalizzabile se e solo se esiste una base di $\mathbb{R}^n$ fatta di autovettori per la A (abbiamo dato queste definizioni lavorando con gli endomorfismi semplici, ma la definizione è analoga per le matrici).

In questo caso avremo che $A = PDP^{-1}$, cioè che diviene diagonale attraverso il cambio di base definito dalla P (o dalla P^{-1}).

Quando lavoriamo in $\mathbb{E}^n$, cioè consideriamo la struttura di spazio vettoriale euclideo standard su $\mathbb{R}^n$, sarà di particolare interesse determinare quando A si può diagonalizzare utilizzando una matrice Q ortonormale, avendo così $A = Q \cdot A \cdot {}^tQ$ (poiché avremo $Q^{-1} = {}^tQ$). In questo caso il cambio di base definito da Q conserverà (per quanto visto precedentemente) il prodotto scalare in $\mathbb{R}^n$, e quindi distanze, angoli, ecc.

Definizione 8.5.1. Una matrice $A \in \mathbb{R}^{n,n}$ si dice *ortogonalmente diagonalizzabile* quando esistono una matrice diagonale D e una matrice ortogonale Q in $\mathbb{R}^{n,n}$ tali che $A = Q \cdot D \cdot {}^tQ$.

Le matrici ortogonalmente diagonalizzabili sono particolarmente interessanti fra tutte quelle diagonalizzabili, come si è detto, perché la trasformazione definita da Q che le diagonalizza conserva la geometria di $\mathbb{R}^n$. Lo scopo di questa sezione è dimostrare un importante teorema (detto appunto *teorema spettrale*) che caratterizza tali matrici, dimostrando che sono tutte e sole le matrici *simmetriche*, cioè quelle per cui ${}^tA = A$.

Osserviamo che se A è ortogonalmente diagonalizzabile, e quindi si ha $A = Q \cdot D \cdot {}^tQ$, allora ciò significa che le colonne di Q (che formano una base ortonormale di $\mathbb{E}^n$), sono anche una base di autovettori per A; infatti se $D = D_{\lambda_1,\ldots,\lambda_n}$, avremo $QD = AQ$, il che significa, se $q_1, \ldots, q_n$ sono le colonne di Q, che:

$$Aq_i = \lambda_i q_i \quad \forall i = 1, \ldots, n.$$

Osserviamo che quando A è una matrice ortogonalmente diagonalizzabile, allora A è *simmetrica*, infatti si ha (sfruttando il fatto che per ogni prodotto di matrici si ha ${}^t(AB) = {}^tB\,{}^tA$, e che ${}^tD = D$):

$${}^tA = {}^t(QD\,{}^tQ) = {}^t({}^tQ)\,{}^tD\,{}^tQ = QD\,{}^tQ = A.$$

Quello a cui vogliamo arrivare (e che costituirà il contenuto del cosiddetto teorema spettrale) è che vale anche il viceversa, cioè che ogni matrice ortogonalmente diagonalizzabile è anche simmetrica. Un primo risultato che ci serve a questo scopo è il seguente:

Proposizione 8.5.1. *Se $A \in \mathbb{R}^{n,n}$ è una matrice simmetrica, allora A ha solo autovalori reali.*

Dimostrazione. Nota Bene: La completa comprensione di questa dimostrazione richiede nozioni sui numeri complessi che possono essere trovate nell'Appendice 12.

Per il teorema fondamentale dell'algebra l'equazione $p_A(t) = 0$, che tramite il polinomio caratteristico individua gli autovalori di A, ha sempre n radici complesse (contate con la propria molteplicità), vogliamo vedere che esse sono tutte reali. Sia quindi λ una soluzione di $p_A(t) = 0$, con $\lambda \in \mathbb{C}$; vogliamo vedere che in effetti $\lambda \in \mathbb{R}$. Se λ è un autovalore complesso, ciò implica che esiste un autovettore $\mathbf{v} \in \mathbb{C}^n$, $\mathbf{v} \neq \mathbf{0}$, tale che $A\mathbf{v} = \lambda\mathbf{v}$.

Consideriamo la seguente operazione $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ (detta prodotto hermitiano), ricordando che il coniugato di un numero complesso $z = a + bi$ è $\overline{z} = a - bi$:

$$\mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^{n} v_i \overline{w_i}.$$

Si può dimostrare facilmente che per ogni matrice $A \in \mathbb{C}^{n,n}$, si ha:

$$\mathbf{v} \cdot A\mathbf{w} = {}^t\overline{(A\mathbf{w})}\mathbf{v}.$$

Consideriamo allora i prodotti (ricordiamo che il modulo di un numero complesso z è dato dal numero reale $|z| = z\overline{z}$):

$$A\mathbf{v} \cdot \mathbf{v} = \lambda\mathbf{v} \cdot \mathbf{v} = \sum_{i=1}^{n} \lambda v_i \overline{v_i} = \lambda \sum_{i=1}^{n} v_i \overline{v_i} = \lambda \sum_{i=1}^{n} |v_i|^2.$$

$$\mathbf{v} \cdot A\mathbf{v} = \mathbf{v} \cdot \lambda\mathbf{v} = \sum_{i=1}^{n} \overline{\lambda v_i} v_i = \overline{\lambda} \sum_{i=1}^{n} \overline{v_i} v_i = \overline{\lambda} \sum_{i=1}^{n} |v_i|^2.$$

Ma poiché A è una matrice simmetrica a elementi reali, avremo

$$\mathbf{v} \cdot A\mathbf{v} = {}^t\overline{(A\mathbf{v})}\mathbf{v} = \overline{{}^t\mathbf{v}\,{}^tA}\mathbf{v} = A\mathbf{v} \cdot \mathbf{v}.$$

Quindi si ottiene che $\lambda(\sum_{i=1}^{n} |v_i|^2) = \overline{\lambda}(\sum_{i=1}^{n} |v_i|^2)$, e poiché $\mathbf{v} \neq \mathbf{0}$, ciò ci dà: $\lambda = \overline{\lambda}$, e quindi $\lambda \in \mathbb{R}$. □

Anche la prossima proposizione costituisce un ulteriore passo verso la dimostrazione del teorema spettrale.

Proposizione 8.5.2. *Sia $A \in \mathbb{R}^{n,n}$ una matrice simmetrica. Siano $\lambda_1, \lambda_2 \in \mathbb{R}$ due autovalori distinti di A, e siano $\mathbf{v}_1 \in V_{\lambda_1}$ e $\mathbf{v}_2 \in V_{\lambda_2}$ due autovettori relativi agli autovalori considerati. Allora $\mathbf{v_1} \cdot \mathbf{v}_2 = 0$, cioè $\mathbf{v}_1$ e $\mathbf{v}_2$ sono ortogonali in $\mathbb{E}^n$.*

Dimostrazione. Notiamo che in $\mathbb{E}^n$ il prodotto scalare $\mathbf{v} \cdot \mathbf{w}$ può essere scritto, quando si considerino i vettori come matrici colonna in $\mathbb{R}^{n,1}$: $\mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^{n} v_i w_i = {}^t\mathbf{v} \cdot \mathbf{w}$. Consideriamo adesso i prodotti:

$$(A\mathbf{v}_1) \cdot \mathbf{v}_2 = (\lambda_1\mathbf{v}_1) \cdot \mathbf{v}_2 = \lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2),$$

$$\mathbf{v}_1 \cdot (A\mathbf{v}_2) = \mathbf{v}_1 \cdot (\lambda_2\mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2).$$

Poiché però si ha anche (in notazione matriciale) che, $\forall\, \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^{n,n}$:

$$(A\mathbf{v}_1) \cdot \mathbf{v}_2 = {}^t(A\mathbf{v}_1)\mathbf{v}_2 = {}^t(\mathbf{v}_1)\,{}^tA\mathbf{v}_2 = {}^t\mathbf{v}_1 A\mathbf{v}_2 = \mathbf{v}_1 \cdot (A\mathbf{v}_2);$$

ne consegue che $\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2)$, e poiché $\lambda_1 \neq \lambda_2$, otteniamo $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. □

Abbiamo adesso tutti gli ingredienti necessari a dimostrare il teorema principale di questa sezione.

Teorema 8.5.1 (Teorema spettrale)**.** *Una matrice $A \in \mathbb{R}^{n,n}$ risulta ortogonalmente diagonalizzabile se e solo se è simmetrica.*

Dimostrazione. Abbiamo già visto che ogni matrice ortogonalmente diagonalizzabile è simmetrica, quindi dobbiamo dimostrare il viceversa: *se A è simmetrica, allora è* anche ortogonalmente diagonalizzabile. Dimostreremo il teorema per induzione su n. Se $n = 1$, $A = (a_{11})$ (ogni matrice in $\mathbb{R}^{1,1}$ è simmetrica); ha 1 come autovettore, $A \cdot 1 = a_{11}.1$, e $\{1\}$ è una base ortonormale per $\mathbb{R}^{1,1} \cong \mathbb{R}$, quindi ammette una base ortonormale di autovettori. Supponiamo ora $n \geq 2$ e che l'enunciato sia vero per $n-1$ (cioè che le matrici simmetriche in $\mathbb{R}^{n-1,n-1}$ siano ortogonalmente diagonalizzabili). Dobbiamo vedere che ciò implica che ogni matrice simmetrica $A \in \mathbb{R}^{n,n}$ è ortogonalmente diagonalizzabile.

Per il teorema fondamentale dell'algebra (cfr. 10.4.3) $p_A(t) = 0$ ha almeno una soluzione complessa λ che per la Proposizione 8.5.1 deve essere reale. Esiste quindi un autovettore $\mathbf{v} \in \mathbb{R}^n$ relativo a λ, e possiamo supporre che $|\mathbf{v}| = 1$ (perché altrimenti basta considerare $\mathbf{v}' = \mathbf{v}/|\mathbf{v}|$). Possiamo completare $\{\mathbf{v}\}$ a una base ortonormale di $\mathbb{E}^n$: $\mathcal{B} = \{\mathbf{v}, \mathbf{v}_2, \dots \mathbf{v}_n\}$ (per il Teorema 8.3.1). Sia P la matrice ortogonale che ha $\mathbf{v}, \mathbf{v}_2, \dots, \mathbf{v}_n$ come sue colonne. Allora la matrice $M_{\mathcal{B},\mathcal{B}}(f_A)$, che esprime rispetto alla base $\mathcal{B}$ l'endomorfismo $f_A : \mathbb{R}^n \to \mathbb{R}^n$, definito dalla matrice A, sarà $M_{\mathcal{B},\mathcal{B}}(f_A) = P^{-1}AP = {}^tPAP$ (in quanto P è la matrice del cambiamento di base dalla base canonica a $\mathcal{B}$).

In particolare, la prima colonna di $M_{\mathcal{B},\mathcal{B}}(f_A)$ è data dal vettore $f_A(\mathbf{v})$, espresso rispetto alle base $\mathcal{B}$: $f_A(\mathbf{v}) = b_{11}\mathbf{v} + b_{21}\mathbf{v}_2 + \cdots + b_{n1}\mathbf{v}_n$; ma poiché $f_A(\mathbf{v}) = A\mathbf{v} = \lambda\mathbf{v}$, avremo: $b_{11} = \lambda$, e $b_{21} = \cdots = b_{n1} = 0$. Notiamo inoltre che la matrice $M_{\mathcal{B},\mathcal{B}}(f_A)$ è simmetrica, in quanto:

$${}^tM_{\mathcal{B},\mathcal{B}}(f_A) = {}^t({}^tPAP) = {}^tP{}^tAP = {}^tPAP = M_{\mathcal{B},\mathcal{B}}(f_A).$$

Quindi $M_{\mathcal{B},\mathcal{B}}(f_A)$ si può scrivere:

$$M_{\mathcal{B},\mathcal{B}}(f_A) = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix},$$

dove B è una matrice simmetrica (perché $M_{\mathcal{B},\mathcal{B}}(f_A)$ è simmetrica). Ora, per ipotesi di induzione, la B è ortogonalmente diagonalizzabile, quindi esiste una matrice ortogonale $Q' \in \mathbb{R}^{n-1,n-1}$, tale che ${}^tQ'BQ' = D'$ è diagonale. Consideriamo la matrice

$$R = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & Q' & \\ 0 & & & \end{pmatrix},$$

che è anch'essa ortogonale perché lo è Q'. Avremo:

$${}^tR \cdot M_{\mathcal{B},\mathcal{B}}(f_A) \cdot R = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & {}^tQ' & \\ 0 & & & \end{pmatrix} \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & Q' & \\ 0 & & & \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & {}^tQ'BQ' & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & D' & \\ 0 & & & \end{pmatrix} = D.$$

Notiamo che D è diagonale poiché D' lo è. Sia ora $Q = PR$; in quanto prodotto di matrici ortogonali Q è ortogonale (infatti ${}^tQQ = {}^t(PR)PR = {}^tR{}^tPPR = I$), e si ha:

$${}^tQAQ = {}^tR{}^tPAPR = {}^tRM_{\mathcal{B},\mathcal{B}}(f_A)R = D.$$

E questo prova che A è ortogonalmente diagonalizzabile, completando la dimostrazione. □

8.6 Esercizi

Esercizio 8.6.1. In $\mathbb{E}^2$, si calcoli il prodotto scalare tra i seguenti vettori, a due a due:

$$\mathbf{v}_1 = (1,0),\ \mathbf{v}_2 = (3,-1),\ \mathbf{v}_3 = (0,4),\ \mathbf{v}_4 = (-10,4),\ \mathbf{v}_5 = (2,6).$$

Quali tra i vettori sono ortogonali tra loro?

Esercizio 8.6.2. In $\mathbb{E}^2$, si calcoli la norma dei seguenti vettori:

$$\mathbf{v}_1 = (1,0\rangle,\ \mathbf{v}_2 = (2-2\rangle,\ \mathbf{v}_3 = (1,-3\rangle,\ \mathbf{v}_4 = (-1,5\rangle,\ \mathbf{v}_5 = (3,7).$$

Esercizio 8.6.3. In $\mathbb{E}^2$, si calcoli la distanza tra le coppie di vettori $(\mathbf{v}_1, \mathbf{v}_2)$ e $(\mathbf{v}_3, \mathbf{v}_4)$ dell'esercizio precedente.

Esercizio 8.6.4. In $\mathbb{E}^3$, si calcoli la proiezione ortogonale $\mathbf{v}_\mathbf{w}$, del vettore $\mathbf{v} = (1,3,2)$ su $\mathbf{w} = (0,1,-1)$. Si determini poi l'angolo tra i due vettori.

Esercizio 8.6.5. In $\mathbb{E}^3$, si determinino i vettori ortogonali a $\mathbf{v} = (1,0,0)$ (cioè lo spazio $\mathbf{v}^\perp$) e quelli di norma 1.

Esercizio 8.6.6. In $\mathbb{E}^3$, si determini il valore del parametro $h \in \mathbb{R}$ tale che i seguenti due vettori siano ortogonali tra loro:

$$\mathbf{v} = (1,h,3), \mathbf{w} = (h,2,-4).$$

Esercizio 8.6.7. In $\mathbb{E}^3$, si determini $h \in \mathbb{R}$ in modo che i seguenti due vettori siano ortogonali tra loro: $\mathbf{v} = (h,0,1), \mathbf{w} = (1,h,-h)$.

Esercizio 8.6.8. In $\mathbb{E}^3$, determinare i vettori ortogonali a $\mathbf{v} = (1,1,0)$ (cioè lo spazio $\mathbf{v}^\perp$). Determinare poi quelli di norma 1.

Esercizio 8.6.9. In $\mathbb{E}^3$, determinare $W^\perp$, ove W sia il sottospazio di $\mathbb{E}^3$ generato da $\{(0,1,0),(1,3,-2)\}$. Determinare poi l'insieme:

$$S = \{(x,y,z) \in W^\perp \mid ||(x,y,z)|| = 1\}.$$

Esercizio 8.6.10. In $\mathbb{E}^3$, determinare $W^\perp$, ove W sia il sottospazio $W \subset \mathbb{E}^3$ generato da $\{(1,1,1),(0,0,-2)\}$. Trovare $\dim W^\perp$ e una sua base.

Esercizio 8.6.11. In $\mathbb{E}^3$, determinare $W^\perp$, ove W sia il sottospazio $W = \{(x,y,z) \in \mathbb{R}^3 \,|\, x - y = 3x = -x + z\} \subset \mathbb{R}^3$, e determinarne la dimensione e una base.

Esercizio 8.6.12. In $\mathbb{E}^3$, sia data la base $\mathcal{B} = \{(0,1,1),(1,0,1),(1,1,0)\}$. Determinare una base ortonormale $\mathcal{B}'$ di $\mathbb{E}^3$ tramite il processo di ortonormalizzazione di Gram-Schmidt.

Esercizio 8.6.13. Si ripeta l'esercizio precedente partendo da

$$\mathcal{B} = \{(1,2,4),(5,-2,1),(0,1,1)\},$$

e poi da

$$\mathcal{B} = \{(0,-2,2),(1,2,1),(1,1,1)\}.$$

Esercizio 8.6.14. In $\mathbb{E}^3$, sia dato il sottospazio $W = \{(x, y, z) \in \mathbb{R}^3 \mid x = 2y = x - z\}$. Trovare una base di W e una di $W^\perp$.

Esercizio 8.6.15. Disegnare W e $W^\perp$ dell'esercizio precedente.

Esercizio 8.6.16. In $\mathbb{E}^4$, sia dato il sottospazio $W = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y = z\}$. Trovare una base di W e una di $W^\perp$.

Esercizio 8.6.17. Calcolare e disegnare nello spazio tridimensionale $U = \{(x, y, z, t) \in \mathbb{R}^4 \mid t = 0\}$ i sottospazi $W \cap U$ e $W^\perp \cap U$ dove W e $W^\perp$ sono stati calcolati nell'esercizio precedente.

Esercizio 8.6.18. Si considerino gli Esercizi 8.6.14, 8.6.16, precedenti. Trovare basi ortogonali per W e $W^\perp$ in ognuno dei tre casi.

Esercizio 8.6.19. In $\mathbb{E}^4$, sia dato il sottospazio $W = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + z = 0, x - t = 0, t + z = 0\}$. Trovare una base di W e una di $W^\perp$.

Esercizio 8.6.20. In $\mathbb{E}^5$, sia dato il sottospazio $W = \{(x, y, z, u, v) \in \mathbb{R}^5 \mid x + z - t = 0, y + z + u = 0, x + 2y + 3z + u = 0, x - v = 0\}$. Trovare una base di W e una di $W^\perp$.

Esercizio 8.6.21. * In $\mathbb{E}^3$, determinare il parametro $h \in \mathbb{R}$ tale che la proiezione ortogonale di $\mathbf{v} = (1, 1, 0)$ sul vettore $\mathbf{w} = (2, 0, 1)$ abbia norma 2.

Esercizio 8.6.22. In $\mathbb{E}^4$, siano: $\mathbf{v}_1 = (1, 0, 1, 0)$, $\mathbf{v}_2 = (3, 1, -1, h)$. Si determini il valore del parametro $h \in \mathbb{R}$ in modo tale che l'angolo tra $\mathbf{v}_1$ e $\mathbf{v}_2$ sia di 45 gradi. Per quale h invece i due vettori sono ortogonali?

Esercizio 8.6.23. In $\mathbb{E}^4$, siano $\mathbf{v}_1 = (1, -1, 1, 1)$, $\mathbf{v}_2 = (0, h, 3, -h)$ con h parametro reale. Si determini il valore di h in modo tale che l'angolo tra $\mathbf{v}_1$ e $\mathbf{v}_2$ sia di 60 gradi. Per quale h invece i due vettori sono ortogonali?

Esercizio 8.6.24. Stabilire se il seguente prodotto su $\mathbb{R}^3$ sia un prodotto scalare: $(x, y, z) \cdot (x', y', z') = 2xx' + xy' + yx' + 5yy' + zz'$. In caso affermativo determinare il sottospazio $\mathbf{v}^\perp$, ove $\mathbf{v} = (1, 0, 0)$ (rispetto a tale prodotto scalare).

Esercizio 8.6.25. * Siano (x, y, z), $(a, b, c) \in \mathbb{R}^3$. Stabilire se le seguenti operazioni $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ sono prodotti scalari:

1. $(x, y, z) \cdot (a, b, c) = ax + 2by - zc$;
2. $(x, y, z) \cdot (a, b, c) = bx + ay$;
3. $(x, y, z) \cdot (a, b, c) = cx + ay + bz$;
4. $(x, y, z) \cdot (a, b, c) = 2ax + 3by + cz$.

Esercizio 8.6.26. In $\mathbb{E}^2$, dire quali dei tre seguenti endomorfismi siano isometrie:
$f(x, y) = (2x + y, x + y)$; $f(x, y) = (y, x)$; $f(x, y) = \left(\frac{\sqrt{2}}{2}x, -\frac{\sqrt{2}}{2}y\right)$.

Esercizio 8.6.27. Stabilire se esiste un valore di $h \in \mathbb{R}$ tale che il seguente endomorfismo di $\mathbb{E}^3$ sia una isometria: $f(x, y, z) = (hy, -hx, hz)$.

Esercizio 8.6.28. Si consideri la seguente applicazione lineare:

$$f(x, y, z) = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ con } A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Tale f è un'isometria di $\mathbb{E}^3$?

Esercizio 8.6.29. Si consideri la seguente applicazione lineare:

$$f(x, y, z) = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ con } A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}.$$

Tale f è un'isometria di $\mathbb{E}^3$?

Esercizio 8.6.30. Determinare, se possibile, dei valori di $h, k \in \mathbb{R}$ tali che la seguente applicazione lineare sia un'isometria: $f : \mathbb{E}^3 \to \mathbb{E}^3$, $f(x, y, z) = (x + 2ky, hz + y, z)$.

Esercizio 8.6.31. Trovare gli $h \in \mathbb{R}$ affinché la seguente applicazione sia una isometria $f : \mathbb{E}^2 \to \mathbb{E}^2$, $f(x, y) = (1/2x, -hy, hx, 1/2y)$. Per tali h disegnare l'immagine della base canonica sotto tale isometria.

Esercizio 8.6.32. * Una rotazione in $\mathbb{E}^3$ è un endomorfismo la cui matrice associata rispetto alla base canonica è la seguente:

$$\begin{pmatrix} \cos\alpha\cos\gamma - \sin\alpha\sin\beta\sin\gamma & -\sin\alpha\cos\beta & -\cos\alpha\sin\gamma - \sin\alpha\sin\beta\cos\gamma \\ \cos\alpha\sin\beta\sin\gamma + \sin\alpha\cos\gamma & \cos\alpha\cos\beta & \cos\alpha\sin\beta\cos\gamma - \sin\alpha\sin\gamma \\ \cos\beta\sin\gamma & -\sin\beta & \cos\beta\cos\gamma \end{pmatrix}.$$

Mostrare che una rotazione in $\mathbb{E}^3$ è un'isometria.

Esercizio 8.6.33. Dire se la seguente applicazione lineare sia un'isometria: $f : \mathbb{E}^n \to \mathbb{E}^n$, $f(x_1, \ldots, x_n) = (x_n, x_{n-1}, \ldots, x_1)$.

Esercizio 8.6.34. Dire, giustificando la risposta, se esiste un sottospazio $W \subset \mathbb{E}^4$ tale che $\dim(W \cap W^\perp) = 1$.

Esercizio 8.6.35. Considerare $(V; \cdot)$, ove $V = \{\mathbf{0}\}$ e $\mathbf{0} \cdot \mathbf{0} = 0 \in \mathbb{R}$. Dire se sia uno spazio vettoriale euclideo. In caso affermativo, dire se se ne può trovare una base ortonormale.

Esercizio 8.6.36. Considerare lo spazio vettoriale euclideo $(\mathbb{R}^2; \cdot)$, ove $(x, y) \cdot (a, b) = xa + 2yb \in \mathbb{R}$. Determinarne una base ortonormale.

Esercizio 8.6.37. Mostrare che il seguente è un prodotto scalare in $\mathbb{R}^2$: $(x, y) \cdot (a, b) = 4xa + 9yb \in \mathbb{R}$. Disegnare nel piano due vettori ortogonali rispetto a tale prodotto scalare.

Esercizio 8.6.38. Sia $A = \begin{pmatrix} 1 & 1 & h \\ 1 & \mathbf{0} & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Trovare h affinché A sia ortogonalmente diagonalizzabile e per tale h calcolare la matrice P che diagonalizza A.

Esercizio 8.6.39. Trovare una ortonormale di autovettori della seguente applicazione lineare (se possibile):

$$f : \mathbb{E}^3 \to \mathbb{E}^3$$

$$(x, y, z) \mapsto (x, y - z, -y + 3z).$$

Esercizio 8.6.40. Trovare se possibile $h \in \mathbb{R}$ tale che la seguente applicazione lineare ammetta una base ortonormale di autovettori:

$$f : \mathbb{E}^3 \to \mathbb{E}^3$$

$$(x, y, z) \mapsto (x - y, h, z).$$

Scrivere poi la base di autovettori.

CAPITOLO 9

Geometria analitica nello spazio

Prerequisiti: Rudimenti di geometria analitica nel piano; sistemi lineari; teorema di Rouché–Capelli; spazi vettoriali.

Obiettivi: Capacità di trattare problemi geometrici nello spazio tridimensionale come posizioni reciproche di rette e piani, loro parallelismo o perpendicolarità.

9.1 Introduzione

9.1.1 Premesse

Lavoreremo nello spazio tridimensionale della geometria euclidea, dove utilizzeremo un sistema di coordinate cartesiane ortogonali. Partiamo dal fissare un punto O nello spazio, e consideriamo l'insieme di tutti i vettori dello spazio applicati in O. Come abbiamo già visto, tale insieme, denotato con il simbolo $\mathcal{V}_O^3$, è uno dei prototipi di spazio vettoriale: in $\mathcal{V}_O^3$ le operazioni di addizione tra vettori e moltiplicazione di questi per gli scalari, sono dotate di tutte quelle proprietà (vedi Capitolo 2) che abbiamo usato per definire la struttura algebrica di *spazio vettoriale*. Inoltre, $\mathcal{V}_O^3$ è stato anche un prototipo (vedi il capitolo precedente) per la definizione della struttura di *spazio euclideo*.

Rivediamo ora come $\mathcal{V}_O^3$ possa essere identificato, tramite un isomorfismo, con lo spazio vettoriale $\mathbb{R}^3$; tale isomorfismo è anche quello che ci consente (come visto nello scorso capitolo), di "trasportare" su $\mathbb{R}^3$ l'operazione di prodotto scalare fra i vettori di $\mathcal{V}_O^3$, rendendo anche $\mathbb{R}^3$ uno spazio euclideo.

Introduciamo un sistema di riferimento cartesiano ortogonale $Oxyz$ in $\mathcal{V}_O^3$; ciò equivale a fissare in $\mathcal{V}_O^3$ tre vettori ortogonali $\left(\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}\right)$ di modulo 1 (cioè $\left\{\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}\right\}$ formano una *base ortonormale* per $\mathcal{V}_O^3$ (vedi

Capitolo 8). Ne segue che ogni vettore $\overrightarrow{OP} \in \mathcal{V}_O^3$ si scrive in modo unico, nella forma:

$$\overrightarrow{OP} = x\,\vec{i} + y\,\vec{j} + z\,\vec{k}$$

dove x, y, z appartengono ad $\mathbb{R}$ e prendono il nome di coordinate del punto P (e, per estensione, anche del vettore $\overrightarrow{OP}$) rispetto alla base $\left\{\vec{i}, \vec{j}, \vec{k}\right\}$ di $\mathcal{V}_O^3$. Resta così definita un'applicazione $j : \mathcal{V}_O^3 \to \mathbb{R}^3$ nel modo seguente:

Sia $\overrightarrow{OP} = x\,\vec{i} + y\,\vec{j} + z\,\vec{k}$, allora $j\left(\overrightarrow{OP}\right) = (x, y, z)$.

Vale inoltre il seguente risultato fondamentale:

Teorema 9.1.1. *L'applicazione $j : \mathcal{V}_O^3 \to \mathbb{R}^3$ sopra definita gode delle seguenti proprietà:*

1. *j è biettiva;*
2. $j\left(\overrightarrow{OP}_1 + \overrightarrow{OP}_2\right) = j\left(\overrightarrow{OP}_1\right) + j\left(\overrightarrow{OP}_2\right), \forall\, \overrightarrow{OP}_1, \overrightarrow{OP}_2 \in \mathcal{V}_O^3$;
3. $j\left(\lambda\overrightarrow{OP}\right) = \lambda j\left(\overrightarrow{OP}\right), \forall\, \overrightarrow{OP} \in \mathcal{V}_O^3, \forall \lambda \in \mathbb{R}$.

Quindi j è un isomorfismo fra i due spazi vettoriali.

L'applicazione j ci permette di identificare $\mathcal{V}_O^3$ e $\mathbb{R}^3$ con le rispettive strutture e quindi di passare dal linguaggio dei vettori a quello dei punti individuati dalla loro terna di coordinate. In questo capitolo useremo spesso il termine "spazio" o "spazio tridimensionale" per indicare indifferentemente $\mathcal{V}_O^3$ o $\mathbb{R}^3$, intendendo anche che passiamo da considerare oggetti nell'uno o nell'altro spazio tramite l'isomorfismo j.

Osservazione 9.1.1. Queste stesse considerazioni possono essere fatte in maniera del tutto analoga considerando sia $\mathcal{V}_O^2$ e $\mathbb{R}^2$ che $\mathcal{V}_O^1$ e $\mathbb{R}^1$.

Prima di passare allo studio degli oggetti geometrici (principalmente rette e piani) nello spazio tridimensionale, ricordiamo la classica definizione di prodotto scalare di due vettori in $\mathcal{V}_O^3$ (esprimendoli poi tramite la base ortonormale $\left\{\vec{i}, \vec{j}, \vec{k}\right\}$):

Definizione 9.1.1. Dati due vettori $\overrightarrow{v_1}, \overrightarrow{v_2} \in \mathcal{V}_O^3$ si dice *prodotto scalare* di $\overrightarrow{v_1}, \overrightarrow{v_2}$ e si indica con $\overrightarrow{v_1} \cdot \overrightarrow{v_2}$, il numero reale

$$\overrightarrow{v_1} \cdot \overrightarrow{v_2} = |\overrightarrow{v_1}||\overrightarrow{v_2}| \cos\alpha,$$

dove α è l'angolo convesso formato dai vettori $\overrightarrow{v_1}, \overrightarrow{v_2}$ e $|\overrightarrow{v_1}|, |\overrightarrow{v_2}|$ denotano, rispettivamente, i moduli dei vettori $\overrightarrow{v_1}, \overrightarrow{v_2}$.

Osservazione 9.1.2. Il prodotto scalare di due vettori è nullo se e solo se essi sono ortogonali, infatti: $\overrightarrow{\mathbf{v_1}} \cdot \overrightarrow{\mathbf{v_2}} = 0$ se e solo se $\alpha = \pi/2$, oppure uno dei due vettori è nullo.
Ricordiamo che il vettore nullo si considera ortogonale ad ogni altro vettore (anche a se stesso) e che per il prodotto scalare di due vettori $\overrightarrow{\mathbf{v_1}}, \overrightarrow{\mathbf{v_2}} \in \mathcal{V}^3_O$, valgono le seguenti proprietà (vedi capitolo precedente):

1. $\overrightarrow{\mathbf{v_1}} \cdot \overrightarrow{\mathbf{v_2}} = \overrightarrow{\mathbf{v_2}} \cdot \overrightarrow{\mathbf{v_1}}$;
2. $\overrightarrow{\mathbf{v_1}} \cdot (\overrightarrow{\mathbf{v_2}} + \overrightarrow{\mathbf{v_3}}) = \overrightarrow{\mathbf{v_1}} \cdot \overrightarrow{\mathbf{v_2}} + \overrightarrow{\mathbf{v_1}} \cdot \overrightarrow{\mathbf{v_3}}$, $\forall\, \overrightarrow{\mathbf{v_3}} \in \mathcal{V}^3_O$;
3. $(\lambda \overrightarrow{\mathbf{v_1}}) \cdot \overrightarrow{\mathbf{v_2}} = \lambda(\overrightarrow{\mathbf{v_1}} \cdot \overrightarrow{\mathbf{v_2}}) = \overrightarrow{\mathbf{v_1}} \cdot (\lambda \overrightarrow{\mathbf{v_2}})$, $\forall\, \lambda \in \mathbb{R}$;
4. il prodotto scalare di un vettore con se stesso dà $\overrightarrow{\mathbf{v_1}} \cdot \overrightarrow{\mathbf{v_1}} = |\overrightarrow{\mathbf{v_1}}|^2$; quindi è sempre positivo ed è nullo se e solo se $\overrightarrow{\mathbf{v_1}}$ è il vettore nullo.

Infine vediamo come il prodotto scalare di due vettori si "rilegga" in $\mathbb{R}^3$ grazie all'identificazione tra $\mathcal{V}^3_O$ e $\mathbb{R}^3$ che abbiamo visto sopra.

Siano $\overrightarrow{\mathbf{OP}} = a\,\overrightarrow{\mathbf{i}} + b\,\overrightarrow{\mathbf{j}} + c\,\overrightarrow{\mathbf{k}}$ e $\overrightarrow{\mathbf{OQ}} = a'\,\overrightarrow{\mathbf{i}} + b'\,\overrightarrow{\mathbf{j}} + c'\,\overrightarrow{\mathbf{k}}$ due vettori di $\mathcal{V}^3_O$; avremo allora (ricordando che $\overrightarrow{\mathbf{i}} \cdot \overrightarrow{\mathbf{i}} = \overrightarrow{\mathbf{j}} \cdot \overrightarrow{\mathbf{j}} = \overrightarrow{\mathbf{k}} \cdot \overrightarrow{\mathbf{k}} = 1$ e $\overrightarrow{\mathbf{i}} \cdot \overrightarrow{\mathbf{j}} = \overrightarrow{\mathbf{j}} \cdot \overrightarrow{\mathbf{k}} = \overrightarrow{\mathbf{k}} \cdot \overrightarrow{\mathbf{i}} = 0$):

$$\overrightarrow{\mathbf{OP}} \cdot \overrightarrow{\mathbf{OQ}} = \left(a\,\overrightarrow{\mathbf{i}} + b\,\overrightarrow{\mathbf{j}} + c\,\overrightarrow{\mathbf{k}}\right) \cdot \left(a'\,\overrightarrow{\mathbf{i}} + b'\,\overrightarrow{\mathbf{j}} + c'\,\overrightarrow{\mathbf{k}}\right) = aa' + bb' + cc'.$$

Quindi l'operazione di prodotto scalare si può considerare anche in $\mathbb{R}^3$, definita dalla legge

$$(a, b, c) \cdot (a', b', c') = aa' + bb' + cc',$$

ed è questa che considereremo come *prodotto scalare standard* in $\mathbb{R}^3$, tramite la quale esprimere tutte le questioni di ortogonalità tra rette e piani espressi tramite equazioni nelle coordinate cartesiane.

Si ha inoltre che il valore $d(AB)$ della distanza fra due punti dello spazio $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ sarà espressa dal modulo del vettore $\overrightarrow{\mathbf{AB}}$, e avremo quindi $|\overrightarrow{\mathbf{AB}}| = |\overrightarrow{\mathbf{OB}} - \overrightarrow{\mathbf{OA}}|$ e da qui, passando ad $\mathbb{R}^3$ tramite le coordinate, si otterrà $d(AB) = |(x_A - x_B, y_A - y_B, z_A - z_B)|$, cioè :

$$d(AB) = \sqrt{(x_A - x_B, y_A - y_B, z_A - z_B) \cdot (x_A - x_B, y_A - y_B, z_A - z_B)} =$$
$$= \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}.$$

9.2 Piani nello spazio

9.2.1 Equazione cartesiana ed equazioni parametriche del piano

Vediamo come è sempre possibile individuare un piano nello spazio tridimensionale grazie ad una equazione che caratterizza i suoi punti.

Teorema 9.2.1. *Ogni piano dello spazio si rappresenta mediante un'equazione della forma:*

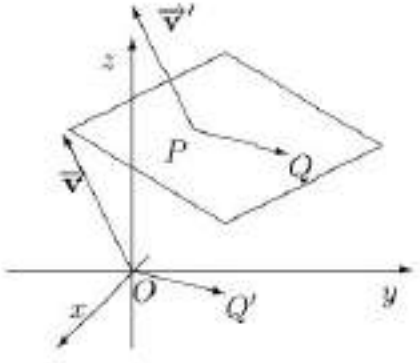
$$ax + by + cz + d = 0$$

con $a, b, c, d \in \mathbb{R}$ e $\vec{\mathbf{v}} = (a, b, c) \neq \vec{0}$ che risulta perpendicolare al piano. Questa prende il nome di equazione cartesiana del piano.

Viceversa, ogni equazione del tipo: $ax + by + cz + d = 0$, dove si abbia $(a, b, c) \neq (0, 0, 0)$, rappresenta un piano nello spazio.

Dimostrazione. Nello spazio tridimensionale, dotato di un sistema di riferimento cartesiano ortogonale, sia dato un piano π; fissiamo un punto $P = (x_0, y_0, z_0)$ sul piano π e consideriamo poi un generico punto $Q = (x, y, z)$ nello spazio. Vogliamo vedere quali condizioni dobbiamo imporre alle coordinate (x, y, z) per avere che Q stia su π.

Sia $\vec{\mathbf{v}} \in \mathcal{V}_O^3$ un vettore tale che il suo parallelo $\vec{\mathbf{v}}'$, applicato in P, sia perpendicolare al piano π; Q starà su π se e solo se $\overrightarrow{PQ} \perp \vec{\mathbf{v}}'$.



Supponiamo che il vettore $\vec{\mathbf{v}}$ abbia coordinate (a, b, c) e consideriamo poi il vettore $\overrightarrow{OQ'}$, parallelo e di uguale modulo a $\overrightarrow{PQ}$; per la regola della somma tra vettori avremo che $\overrightarrow{OQ} = \overrightarrow{OQ'} + \overrightarrow{OP}$, e quindi

$$\overrightarrow{OQ'} = \overrightarrow{OQ} - \overrightarrow{OP} = (x - x_0, y - y_0, z - z_0).$$

Si ha che $\overrightarrow{PQ} \perp \vec{v}'$ se e solo se $\overrightarrow{OQ'} \perp \vec{v}$, e, per quanto visto sul prodotto scalare, avremo: $\overrightarrow{PQ} \perp \vec{v}'$ se e solo se $\overrightarrow{OQ'} \perp \vec{v}$, cioè se e solo se: $(x - x_0, y - y_0, z - z_0) \cdot (a, b, c) = 0$, da cui:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

Per concludere, posto $d = -ax_0 - by_0 - cz_0$, otteremo che le condizioni affinché Q stia sul piano sono date proprio dalla seguente equazione che viene pertanto chiamata *equazione cartesiana* del piano:

$$ax + by + cz + d = 0.$$

Viceversa, avremo che data un'equazione come la precedente, scelto un qualsiasi punto $P = (x_0, y_0, z_0)$ che la soddisfa, essa rappresenterà (tramite la costruzione appena vista) il piano passante per P e perpendicolare al vettore $\vec{v}$ di coordinate (a, b, c). □

Notiamo che ogni altra equazione del tipo: $lax + lby + lcz + ld = 0$, per $l \neq 0$, rappresenta ovviamente lo stesso piano.

Ogni piano nello spazio può essere rappresentato sia in forma "implicita", mediante una equazione cartesiana, come visto sopra, sia in forma "esplicita" attraverso delle equazioni parametriche.

Teorema 9.2.2. *Dato nello spazio $\mathbb{R}^3$ un piano π e un suo punto $P = (x_1, y_1, z_1)$, ogni punto $Q = (x, y, z)$ del piano si può rappresentare mediante equazioni della forma:*

$$\begin{cases} x = x_1 + a_1 u + b_1 t \\ y = y_1 + a_2 u + b_2 t \\ z = z_1 + a_3 u + b_3 t \end{cases}$$

dove $\vec{v}_1 = (a_1, a_2, a_3)$ e $\vec{v}_2 = (b_1, b_2, b_3)$ sono due vettori linearmente indipendenti e paralleli al piano (ciò si esprime anche dicendo che $\{\vec{v}_1, \vec{v}_2\}$ costituisce una base della "giacitura" del piano), mentre u, t sono parametri reali. Queste sono dette equazioni parametriche *del piano.*

Dimostrazione. Nella Figura 9.1 i vettori $\vec{v}_1$ e $\vec{v}_2$ sono paralleli ai vettori $\overrightarrow{PQ_1}$ e $\overrightarrow{PQ_2}$, giacenti sul piano. Ogni vettore $\overrightarrow{PQ}$ che stia sul piano può essere scritto come combinazione lineare di $\overrightarrow{PQ_1}$ e $\overrightarrow{PQ_2}$, vale a dire nella forma: $\overrightarrow{PQ} = u\overrightarrow{PQ_1} + t\overrightarrow{PQ_2}$. Cioè equivalente a:

$$\overrightarrow{OQ'} = u\vec{v}_1 + t\vec{v}_2$$

ossia

$$(x - x_1, y - y_1, z - z_1) = u(a_1, a_2, a_3) + t(b_1, b_2, b_3),$$

da cui si ricavano le equazioni parametriche indicate sopra. □

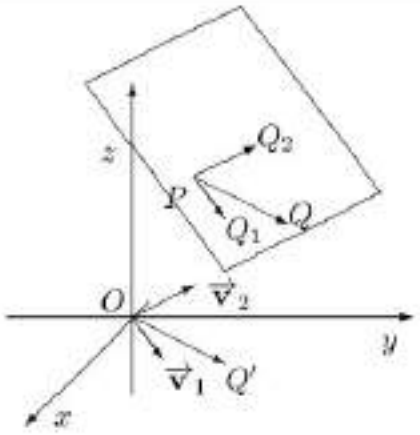


Figura 9.1: Visualizzazione delle equazioni parametriche di un piano.

Piano: Complementi

I seguenti risultati si dimostrano a partire dai teoremi precedenti.

Corollario 9.2.1. *Dati nello spazio tre punti distinti non allineati P_1, P_2, P_3 , l'equazione del piano passante per essi è data da:*

$$\det \begin{pmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{pmatrix} = 0$$

Dimostrazione. Abbiamo visto nella dimostrazione del Teorema 9.2.2 che, perché un punto (x, y, z) stia sul piano, si deve avere:

$$(x-x_1, y-y_1, z-z_1) = u(a_1, a_2, a_3) + t(b_1, b_2, b_3).$$

Ciò equivale a dire che il vettore $(x-x_1, y-y_1, z-z_1)$ non è linearmente indipendente dai vettori

$$\vec{\mathbf{v}}_1 = (a_1, a_2, a_3) \text{ e } \vec{\mathbf{v}}_2 = (b_1, b_2, b_3).$$

Ma nel caso presente (vedi Figura 9.2, ove P_1, P_2, P_3 svolgono il ruolo prima di O, Q_1, Q_2), si ha che:

$$\vec{\mathbf{v}}_1 = (a_1, a_2, a_3) = (x_2-x_1, y_2-y_1, z_2-z_1)$$

e

$$\vec{\mathbf{v}}_2 = (b_1, b_2, b_3) = (x_3-x_1, y_3-y_1, z_3-z_1)$$

e quindi che le tre righe della matrice

$$\det \begin{pmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{pmatrix}$$

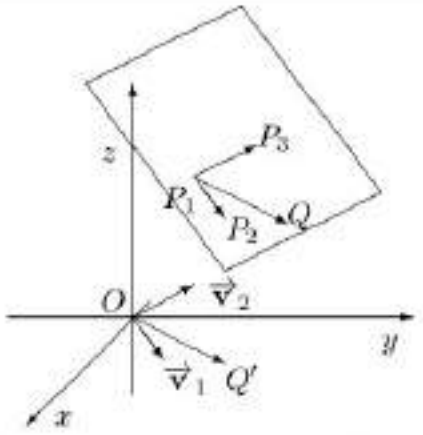


Figura 9.2: Piano passante per Q e parallelo a $P_2 - P_1$ e $P_3 - P_1$.

non devono essere linearmente indipendenti, il che equivale al fatto che il suo determinante sia nullo. □

Corollario 9.2.2. *Ogni piano passante per un punto $P_0 = (x_0, y_0, z_0)$, ha un' equazione cartesiana della forma:*

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0;$$

con $(a, b, c) \neq (0, 0, 0)$.

Corollario 9.2.3. *Un piano passa per l'origine se e solo se ha un'equazione cartesiana del tipo:*

$$ax + by + cz = 0.$$

Piani paralleli

Vediamo adesso quando è che due piani si dicono paralleli e come determinare le condizioni di parallelismo.

Definizione 9.2.1. Due piani π e π' dello spazio $\mathcal{V}_0^3$ si dicono paralleli, se coincidono, oppure se non hanno punti in comune, e si scrive: $\pi \parallel \pi'$. Dunque:

$$\pi \parallel \pi' \text{ se e solo se } \pi = \pi' \text{ oppure } \pi \cap \pi' = \emptyset.$$

Teorema 9.2.3. *Siano dati, nello spazio $\mathcal{V}_O^3$, due piani di equazioni:*

$$\pi : ax + by + cz + d = 0; \quad \pi' : a'x + b'y + c'z + d' = 0;$$

si hanno allora le seguenti due possibilità:

1. *π e π' sono **paralleli** se e solo se:*

$$r\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 1$$

*inoltre π e π' sono **paralleli disgiunti** o **paralleli coincidenti** a seconda che la matrice:*

$$\begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \end{pmatrix}$$

abbia rispettivamente rango 2 oppure 1.

2. *π e π' si intersecano in una retta, e si dicono **incidenti**, se e solo se:*

$$r\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2.$$

Dimostrazione. Entrambe le cose si deducono osservando che il sistema:

$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

nel caso 2 ha infinite soluzioni con una variabile libera, mentre nel caso 1 se

$$r\begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \end{pmatrix} = 2,$$

il sistema non ha soluzioni, mentre ne avrà infinite (con due variabili libere), se

$$r\begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \end{pmatrix} = 1.$$

□

Piani perpendicolari

Consideriamo adesso la condizione di perpendicolarità tra piani nello spazio.

Definizione 9.2.2. Dato un piano π e un punto $S \in \pi$, si dice che un vettore $\overrightarrow{\mathbf{ST}}$ è perpendicolare a π, e si scrive $\overrightarrow{\mathbf{ST}} \perp \pi$, se è perpendicolare a ogni vettore $\overrightarrow{\mathbf{SQ}}$ con $Q \in \pi$.

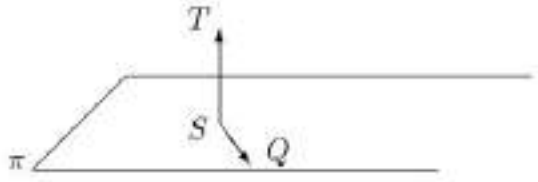


Figura 9.3: Il vettore $\mathbf{ST}$ è perpendicolare al piano π.

Definizione 9.2.3. Due piani π, π' si dicono **perpendicolari**, e si scrive $\pi \perp \pi'$, se si intersecano in una retta r e, comunque fissato un punto $S \in r$, esiste $T \in \pi - r$ tale che

$$\overrightarrow{\mathbf{ST}} \perp \pi'.$$

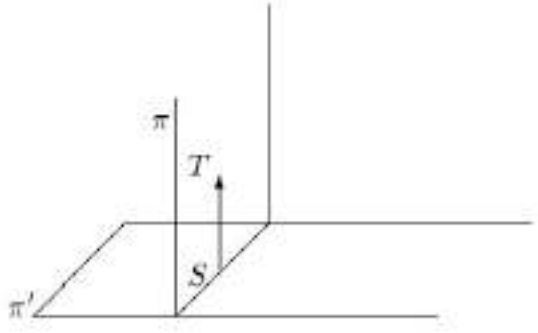


Figura 9.4: Piani π e π' perpendicolari.

Osserviamo che naturalmente, se $\pi \perp \pi'$, allora esiste anche un punto $T' \in \pi' - r$, tale che $\overrightarrow{\mathbf{ST}'} \perp \pi$.

Teorema 9.2.4. *Dati nello spazio due piani di equazioni, rispettivamente:*

$$ax + by + cz + d = 0, \quad a'x + b'y + c'z + d' = 0$$

si ha: $\pi \perp \pi'$ se e solo se $aa' + bb' + cc' = 0$.

Dimostrazione. Ricordiamo che dato un piano nello spazio, di equazione $ax + by + cz + d = 0$, il vettore di $\mathcal{V}_0^3$ definito dal punto di coordinate (a, b, c) è ad esso perpendicolare; allora considerando i vettori di $\mathcal{V}_0^3$ definiti dai punti di coordinate (a, b, c) e (a', b', c'), risulta che $\pi \perp \pi'$ se e solo se $(a, b, c) \perp (a', b', c')$ se e solo se $(a, b, c) \cdot (a', b', c') = 0$. Ciò equivale a:

$$aa' + bb' + cc' = 0.$$

□

9.3 Rette nello spazio

9.3.1 La retta, equazioni cartesiane e parametriche

Osserviamo che ogni retta nello spazio tridimensionale della geometria euclidea può essere definita tramite l'intersezione di due piani distinti che la contengono.

Teorema 9.3.1. *Ogni* ***retta*** *dello spazio tridimensionale dotato di coordiante cartesiane si può rappresentare mediante un sistema di due equazioni della forma:*

$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

tale che

$$rg \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2.$$

Tali equazioni prendono il nome di ***equazioni cartesiane*** *della retta. Viceversa ogni sistema di equazioni come quello sopra rappresenta una retta nello spazio tridimensionale.*

Dimostrazione. Sia data una retta arbitraria r dello spazio tridimensionale e consideriamo due piani distinti π e π' contenenti r, di equazioni:

$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}.$$

Visto che π e π' non sono paralleli, per il Teorema 9.2.3 si ha:

$$rg \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2$$

e quindi, sempre in virtù dello stesso teorema, i punti della retta r sono rappresentati proprio dalle soluzioni del sistema formato dalle due equazioni.

Viceversa dato un sistema di questo tipo, esso rappresenta certamente una retta dello spazio, in virtù del Teorema 9.2.3. □

Equazioni parametriche

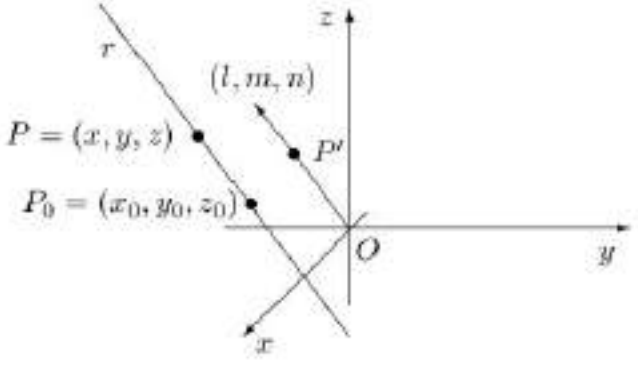
Ogni retta nello spazio può essere rappresentata sia in forma "implicita", mediante equazioni cartesiane, come abbiamo appena visto, che in forma esplicita attraverso delle equazioni parametriche.

Teorema 9.3.2. *Ogni retta dello spazio si può rappresentare mediante equazioni della forma:*

$$\begin{cases} x = lt + x_0 \\ y = mt + y_0 \\ z = nt + z_0 \end{cases}$$

dove $x_0, y_0, z_0, l, m, n \in \mathbb{R}$, $(l, m, n) \neq (0, 0, 0)$ *e* t *è un parametro reale. Queste prendono il nome di* ***equazioni parametriche*** *della retta considerata.*

Viceversa, ogni terna di equazioni di questo tipo rappresenta una retta dello spazio.



Dimostrazione. Data nello spazio una retta r, consideriamo un vettore $\vec{\mathbf{v}} = (l, m, n)$ in $\mathcal{V}_O^3$, parallelo a r. Sia $P_0 = (x_0, y_0, z_0)$ un punto di r e sia invece P un punto generico dello spazio; cerchiamo di esprimere la condizione che $P \in r$.

Avremo che $P \in r$ se e solo se $\overrightarrow{\mathbf{P_0P}}$ è parallelo a $\vec{\mathbf{v}}$; consideriamo allora il vettore $\overrightarrow{\mathbf{OP'}}$ parallelo e di modulo uguale al vettore $\overrightarrow{\mathbf{P_0P}}$: avremo allora che $P \in r$ se e solo se $\overrightarrow{\mathbf{OP'}}$ è parallelo a $\vec{\mathbf{v}}$, e quindi (essendo entrambi applicati nell'origine) se e solo se esiste un numero reale $t \in \mathbb{R}$, per cui $\overrightarrow{\mathbf{OP'}} = t\,\vec{\mathbf{v}}$. Per la regola della differenza tra vettori, si ha:

$$\overrightarrow{\mathbf{OP'}} = \overrightarrow{\mathbf{OP}} - \overrightarrow{\mathbf{OP_0}},$$

che corrisponde, in coordinate, a $(x_1 - x_0, y_1 - y_0, z_1 - z_0)$.

Quindi, in conclusione, si ha che $P \in r$ se e solo se esiste un numero reale t per cui si abbia:

$$P = (x, y, z) = (x_0, y_0, z_0) + t(x_1 - x_0, y_1 - y_0, z_1 - z_0)$$

e cioè

$$x = x_0 + t(x_1 - x_0), \quad y = y_0 + t(y_1 - y_0), \quad z = z_0 + t(z_1 - z_0),$$

da cui la tesi, ponendo:

$$l = x_1 - x_0, \quad m = y_1 - y_0, \quad n = z_1 - z_0.$$

Viceversa, data una terna di equazioni del tipo:

$$\begin{cases} x = lt + x_0 \\ y = mt + y_0 \\ z = nt + z_0 \end{cases}$$

si verifica (in modo analogo a quanto appena visto) che queste sono le equazioni parametriche della retta per P_0, parallela al vettore (l, m, n), e il teorema è dimostrato. □

Osservazione 9.3.1. Notiamo che data una retta r tramite delle equazioni parametriche, ad ogni valore di t corrisponde un punto di r. Se pensiamo a t come la variabile *tempo*, si può visualizzare la r come la traiettoria di un punto nello spazio che si muova di moto rettilineo uniforme con velocità $\vec{\mathbf{v}} = (l, m, n)$ (infatti l, m, n sono le derivate delle funzioni x, y, z rispetto a t) e punto di partenza P_0 (per $t = 0$).

Osservazione 9.3.2. Osserviamo che se l'origine degli assi, O, appartiene alla retta, le equazioni parametriche si possono semplificare: prendendo $P_0 = (0, 0, 0)$, esse diventano

$$\begin{cases} x = lt \\ y = mt \\ z = nt \end{cases}.$$

Notiamo però che non tutte le rette passanti per l'origine saranno necessariamente espresse così come il seguente esempio mostra.

Esempio 9.3.1. La retta di equazioni parametriche:

$$\begin{cases} x = 2t - 2 \\ y = t - 1 \\ z = -t + 1 \end{cases}$$

passa per l'origine (basta assegnare a t il valore $t = 1$).

9.3.2 Retta per due punti dati

Dati arbitrariamente due punti distinti P_1, P_0 nello spazio, vogliamo determinare la (unica) retta r passante per i due punti. Vediamo prima la ricerca delle equazioni parametriche di r.

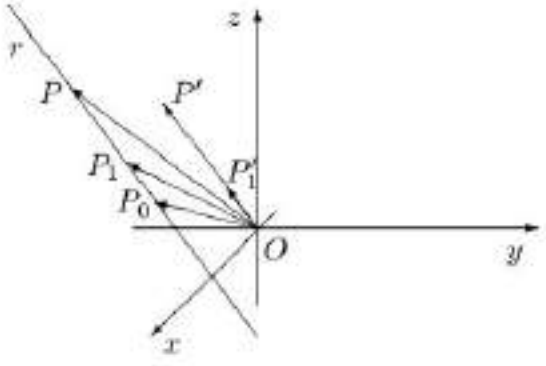
Proposizione 9.3.1. *Siano P_1, P_0 due punti dello spazio, di coordinate $P_1 = (x_1, y_1, z_1)$, $P_0 = (x_0, y_0, z_0)$; allora la retta r contenente i due punti ha equazioni parametriche:*

$$\begin{cases} x = lt + x_0 \\ y = mt + y_0 \\ z = nt + z_0 \end{cases}$$

ove

$$l = x_1 - x_0, \quad m = y_1 - y_0, \quad n = z_1 - z_0.$$

Dimostrazione. Sia $P = (x, y, z)$ un punto generico dello spazio; cerchiamo di esprimere la condizione che $P \in r$.



Consideriamo i vettori $\overrightarrow{\mathbf{OP}_1'}$ $\overrightarrow{\mathbf{OP}'}$ paralleli rispettivamente ai vettori $\overrightarrow{\mathbf{P_0P_1}}$, $\overrightarrow{\mathbf{P_0P}}$; essi sono dati, per la regola della differenza tra vettori, da:

$$\overrightarrow{\mathbf{OP}_1'} = \overrightarrow{\mathbf{OP}_1} - \overrightarrow{\mathbf{OP}_0}, \quad \overrightarrow{\mathbf{OP}'} = \overrightarrow{\mathbf{OP}} - \overrightarrow{\mathbf{OP}_0}.$$

Poiché $P \in r$ se e solo se esiste un numero reale t per cui si abbia $\overrightarrow{\mathbf{P_0P}} = t\overrightarrow{\mathbf{P_0P_1}}$, ciò equivale a dire, usando le coordinate dei punti:

$$P = (x, y, z) = (x_0, y_0, z_0) + t(x_1 - x_0, y_1 - y_0, z_1 - z_0)$$

e cioè

$$x = x_0 + t(x_1 - x_0), \quad y = y_0 + t(y_1 - y_0), \quad z = z_0 + t(z_1 - z_0),$$

da cui la tesi, ponendo:

$$l = x_1 - x_0, \quad m = y_1 - y_0, \quad n = z_1 - z_0.$$

□

Vediamo ora come trovare invece la retta r per P_1, P_0 usando la sua rappresentazione cartesiana.

Proposizione 9.3.2. *Siano P_1, P_0 due punti dello spazio, di coordinate $P_1 = (x_1, y_1, z_1)$, $P_0 = (x_0, y_0, z_0)$: allora la retta r contenente i due punti ha equazioni cartesiane:*

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0} = \frac{z - z_0}{z_1 - z_0}$$

ove si abbia $x_1 - x_0 \neq 0, y_1 - y_0 \neq 0, z_1 - z_0 \neq 0$. Qualora invece in uno o due di questi casi si abbia un'uguaglianza, alla/e equazione/i che perdono di senso (diventando nullo un denominatore) si sostituiscono le opportune equazioni $x = x_0$, $y = y_0$ o $z = z_0$.

Esempio 9.3.2. Siano $P_0 = (1, 6, 3)$, $P_1 = (-1, 0, 2)$; allora la retta r avrà equazioni:

$$\begin{cases} \frac{x-1}{-1-1} = \frac{y-6}{0-6} \\ \frac{x-1}{-1-1} = \frac{z-3}{2-3} \end{cases} ; \quad \begin{cases} \frac{x-1}{-2} = \frac{y-6}{-6} \\ \frac{x-1}{-2} = \frac{z-3}{-1} \end{cases} ; \quad \begin{cases} 3x - y + 3 = 0 \\ x - 2z + 5 = 0 \end{cases}.$$

Esempio 9.3.3. Siano $P_0 = (1, 2, 3)$, $P_1 = (-1, 2, 2)$; allora la retta r avrà equazioni:

$$\begin{cases} y = 2 \\ \frac{x-1}{-1-1} = \frac{z-3}{2-3} \end{cases} ; \quad \begin{cases} y = 2 \\ x - 2z + 5 = 0 \end{cases}.$$

Esempio 9.3.4. Siano $P_0 = (1, 2, 3)$, $P_1 = (1, 2, 2)$; allora la retta r avrà equazioni:

$$\begin{cases} y = 2 \\ x = 1 \end{cases}.$$

9.3.3 Calcolo dei parametri direttori di una retta

Vediamo concretamente come possono essere calcolati i parametri direttori di una retta r, cioè una terna (l, m, n) che indica la direzione della retta nello spazio (ciò equivale a dire il che vettore $\overrightarrow{\mathbf{v}} = (l, m, n)$, è parallelo ad r). Esaminiamo i diversi casi possibili:

- Se la retta r è definita mediante equazioni parametriche della forma

$$\begin{cases} x = lt + x_0 \\ y = mt + y_0 \\ z = nt + z_0 \end{cases}$$

i parametri direttori (l, m, n) sono forniti esplicitamente dalle equazioni stesse.

Esempio 9.3.5. Consideriamo le rette:

$$\begin{cases} x = 4t + 3 \\ y = -3t \\ z = -t - 3 \end{cases}, \quad \begin{cases} x = 4 \\ y = 3t \\ z = 5 \end{cases}.$$

La prima retta ha parametri direttori $(4, -3, -1)$; la seconda $(0, 3, 0)$.

- Se la retta r è definita come retta congiungente due punti distinti assegnati $P_1 = (x_1, y_1, z_1)$ e $P_2 = (x_2, y_2, z_2) \in r$, il calcolo è immediato ricordando la dimostrazione della Proposizione 9.3.1: $l = x_2 - x_1$, $m = y_2 - y_1$, $n = z_2 - z_1$.

Esempio 9.3.6. Trovare l'equazione parametrica di una retta r passante per i punti $P_1 = (1, -1, 3)$ e $P_2 = (2, -1, 4)$.

Il vettore $(l, m, n) = (x_2 - x_1, y_2 - y_1, z_2 - z_1) = (1, 0, 1)$ sarà il vettore direttore e le equazioni parametriche saranno:

$$\begin{cases} x = t + 1 \\ y = -1 \\ z = t + 3 \end{cases}.$$

- Se la retta r è assegnata mediante equazioni cartesiane della forma:

$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

si può procedere in diversi modi:

1. calcolando due soluzioni distinte del sistema lineare, quindi trovando due punti distinti della retta r e utilizzando poi la regola precedente;
2. Trasformando le equazioni in parametriche e ottenendo da queste direttamente i parametri direttori (vedi prossima sezione);

3. Usufruendo del prodotto vettoriale (vedi Sezione 9.4) con il quale si ottiene:

$$l = \det\begin{pmatrix} b & c \\ b' & c' \end{pmatrix}, \quad m = -\det\begin{pmatrix} a & c \\ a' & c' \end{pmatrix}, \quad n = \det\begin{pmatrix} a & b \\ a' & b' \end{pmatrix}.$$

Esempio 9.3.7. Data una retta r di equazioni cartesiane:

$$\begin{cases} x + y + 5 = 0 \\ x - y + 2z = 0 \end{cases}$$

avremo che

$$l = \det\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} = 2, \quad m = -\det\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} = -2,$$

$$n = \det\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2$$

quindi $(l, m, n) = (2, -2, -1)$.

9.3.4 Passaggio da equazioni parametriche a cartesiane e viceversa

Affrontando gli esercizi ci accorgeremo che a volte sono più utili le equazioni cartesiane della retta e altre volte quelle parametriche, quindi è importante saper passare dalle une alle altre e viceversa. Vediamo i due procedimenti utilizzando qualche semplice esempio:

Da parametriche a cartesiane

Se abbiamo una retta in equazioni parametriche come:

$$\begin{cases} x = t + 1 \\ y = 2t + 2 \\ z = 4t + 1 \end{cases}$$

possiamo ottenere le cartesiane nel seguente modo:

- ricaviamo il parametro t dalle equazioni e considerando le relazioni ottenute:
$$t = (x - 1) = (y - 2)/2 = (z - 1)/4;$$
- ricaviamo le equazioni cartesiane, eliminando il parametro t e uguagliando il primo e secondo membro, e poi il primo e il terzo:
$$\begin{cases} x - 1 = (y - 2)/2 \\ x - 1 = (z - 1)/4 \end{cases} \quad \begin{cases} 2x - y = 0 \\ 4x - z = 3 \end{cases}.$$
Ovviamente avremmo potuto uguagliare il primo al secondo e poi il secondo al terzo ottenendo altre equazioni cartesiane equivalenti alle precedenti.

Nell'eliminare il parametro t l'unica accortezza che bisogna avere è che qualora in una delle equazioni manchi il parametro, tale equazione può essere assunta direttamente come una delle cartesiane.

Da cartesiane a parametriche

Per passare dalle equazioni cartesiane alle parametriche, basta risolvere il sistema lineare formato dalle equazioni cartesiane della retta. Tale sistema possiede sicuramente soluzioni e risolvendo rispetto a una variabile libera usata come parametro avremo delle equazioni parametriche. Ad esempio:

$$\begin{cases} 2x + z - 1 = 0 \\ y - z + 1 = 0 \end{cases}.$$

Risolvendo rispetto alla variabile x otteniamo:

$$\begin{cases} z = -2x + 1 \\ y = -2x + 1 - 1 \end{cases}, \quad \begin{cases} z = -2x + 1 \\ y = -2x \end{cases},$$

quindi, ponendo $x = t$:

$$\begin{cases} x = t \\ y = -2t \\ z = -2t + 1 \end{cases}.$$

9.3.5 Posizioni reciproche di due rette

Definizione 9.3.1. Date nello spazio due rette r, r' si dice che:

- r, r' sono complanari se sono contenute nello stesso piano;
- r è parallela a r' (e si scrive: $r \parallel r'$) se r e r' sono complanari e inoltre, o coincidono o non hanno punti in comune;
- r, r' sono incidenti se hanno un unico punto in comune;
- r, r' sono perpendicolari, e si scrive $r \perp r'$, se i vettori direttori delle due rette sono perpendicolari.
  Possiamo notare che nello spazio esisteranno rette perpendicolari disgiunte, ovvero senza alcun punto in comune e rette perpendicolari incidenti.
- r, r' sono sghembe, se non sono complanari. Notiamo che r, r' sono sghembe se e solo se r e r' non sono parallele e $r \cap r' = \emptyset$.

Nella Figura 9.5, r e r' sono due rette sghembe, una sul piano xy e una su quello yz. Nella Figura 9.6 le due rette r e r' sono parallele ed entrambe sul piano xy.

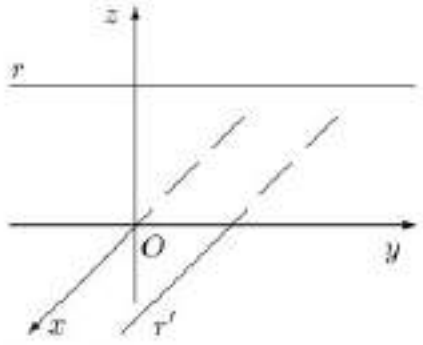


Figura 9.5: Due rette sghembe r e r'.

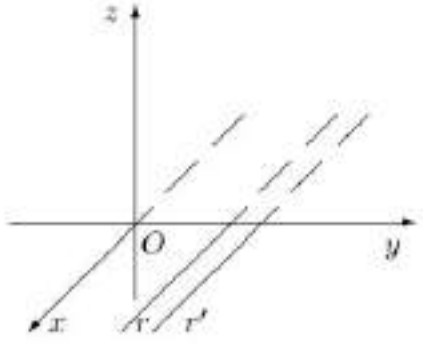


Figura 9.6: Due rette parallele r e r'.

Osservazione 9.3.3. Due rette r e r' sono complanari se e solo se una delle condizioni seguenti è verificata:

1. r e r' sono parallele;
2. r e r' sono incidenti.

In particolare r e r' sono complanari se e solo se non sono sghembe.

Teorema 9.3.3. *Date nello spazio due rette r, r', i cui vettori direttori sono rispettivamente (l, m, n) e (l', m', n') si ha:*

- *r e r' sono parallele se e solo se*

$$r\begin{pmatrix} l & m & n \\ l' & m' & n' \end{pmatrix} = 1;$$

- *r e r' sono perpendicolari se e solo se*

$$ll' + mm' + nn' = 0.$$

Dimostrazione. L'enunciato è una conseguenza immediata del fatto che r, r' sono parallele o perpendicolari se e solo se lo sono i loro vettori direzione. □

9.3.6 Posizioni reciproche di retta e piano

Consideriamo ora tutti i casi che si possono presentare considerando una retta e un piano e la loro posizione relativa.

Definizione 9.3.2. Dati nello spazio una retta r e uno piano π, si dice che π è *parallelo* a r o, equivalentemente, che r è parallela a π, e si scrive $r \parallel \pi$, ossia $\pi \parallel r$, se r è contenuta in π oppure se r e π non hanno punti in comune. In breve:

$$r \parallel \pi \text{ se e solo se } r \subset \pi \text{ oppure } r \cap \pi = \emptyset.$$

Se π e r non sono paralleli, allora si dicono *incidenti*, e hanno uno e un sol punto in comune.

Vediamo come si determina la posizione reciproca di un piano e una retta, se sono assegnate le loro equazioni cartesiane.

Proposizione 9.3.3. *Consideriamo una retta r e un piano π rispettivamente di equazioni cartesiane:*

$$r : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}, \quad \pi : a''x + b''y + c''z + d'' = 0,$$

1. r e π sono paralleli se e solo se

$$r \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = 2 \tag{9.1}$$

Se la (9.1) è soddisfatta allora r e π possono essere paralleli e disgiunti, oppure r essere contenuta in π, a seconda che per la matrice del sistema tra piano e retta si abbia:

$$r \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = 3 \text{ (nessuna soluzione)}$$

oppure

$$r \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = 2 \text{ (infinite soluzioni).}$$

2. *Se r e π non sono paralleli, allora sono incidenti e $r \cap \pi$ consiste di un solo punto; ciò avviene se e solo se:*

$$r \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = 3 \text{ (il sistema ha un'unica soluzione).}$$

La proposizione è conseguenza di quanto noto sui sistemi lineari (teorema di Rouché–Capelli).

Nel caso in cui la retta sia data tramite equazioni parametriche, avremo:

Proposizione 9.3.4. *Consideriamo una retta r e un piano π, dati dalle equazioni:*

$$r : \begin{cases} x = lt + x_0 \\ y = mt + y_0 \\ z = nt + z_0 \end{cases}, \quad \pi : ax + by + cz + d = 0.$$

1. *Avremo che r e π sono incidenti se e solo se:*

$$al + bm + cn \neq 0.$$

In questo caso l'equazione

$$a(lt + x_0) + b(mt + y_0) + c(nt + z_0) + d = 0 \tag{9.2}$$

risulta avere un'unica soluzione in t.

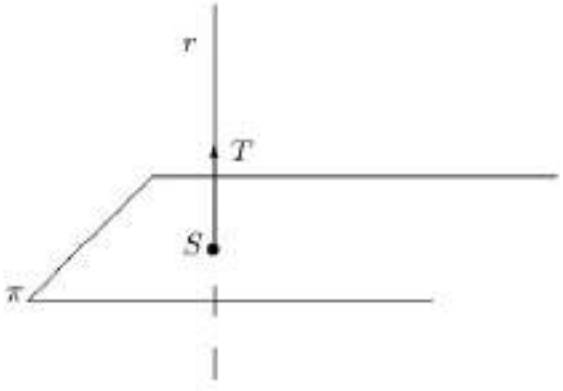
2. *Se invece $al + bm + cn = 0$, allora r e π sono paralleli, e in particolare se l'equazione (9.2) non ha soluzioni in t, allora r e π sono paralleli disgiunti, mentre se (9.2) è un'identità, allora $r \subset \pi$.*

La proposizione segue dal fatto che r e π sono paralleli se e solo se il vettore (a, b, c) associato al piano e il vettore direttore (l, m, n) della retta sono perpendicolari, e quindi se e solo se $(a, b, c) \cdot (l, m, n) = 0$. Il fatto che l'equazione (9.2) non abbia soluzioni ci dice che per nessun valore di t si ottiene un punto della retta che stia su π, quindi $r \cap \pi = \emptyset$, mentre se l'equazione è un'identità , ogni valore di t ne è una soluzione e quindi tutti i punti di r stanno in π.

Retta e piano perpendicolari

Un caso particolare di incidenza si ha quando la retta e il piano sono perpendicolari.

Definizione 9.3.3. Si dice che una retta r è perpendicolare ad un piano π (e si scrive $r \perp \pi$) se la retta e il piano si intersecano in un punto, diciamolo S, e inoltre la retta r contiene un vettore $\overrightarrow{\mathbf{ST}}$ perpendicolare a π con $T \in r$.



Teorema 9.3.4. *Data una retta r il cui vettore direttore sia (l, m, n), e un piano π di equazione:*

$$ax + by + cz + d = 0,$$

si ha

$$r \perp \pi \quad \textit{se e solo se} \quad \mathbf{r}\begin{pmatrix} a & b & c \\ l & m & n \end{pmatrix} = \mathbf{1}.$$

Dimostrazione. Denotiamo con r_0 (e con π_0) la retta (il piano) parallela ad r (parallelo a π) e passante per l'origine e consideriamo i punti:

$$S = (l, m, n), \quad P = (a, b, c),$$

allora:

$r \perp \pi$ se e solo se $r_0 \perp \pi_0$ se e solo se $\overrightarrow{\mathbf{OP}}$ e $\overrightarrow{\mathbf{OS}}$ sono allineati.

Ciò equivale a dire che i vettori (a, b, c) e (l, m, n) , in $\mathbb{R}^3$, sono linearmente dipendenti, da cui la tesi. □

Esercizio 9.3.1. Sia π il piano di equazione $3x - y + 2z + 1 = 0$ e sia A il punto $(1, 2, 3)$. Determinare la retta r perpendicolare a π e passante per A. Poiché il vettore (l, m, n) della retta deve essere dipendente da $(a, b, c) = (3, 2, 1)$, e quindi un suo multiplo, potremo prendere direttamente $(l, m, n) = (3, 2, 1)$, e, poiché r deve passare per A, avremo per r le equazioni parametriche:

$$\begin{cases} x = 3t + 1 \\ y = 2t + 2 \\ z = t + 3 \end{cases}.$$

Un caso in cui si utilizza la perpendicolarità fra piano e retta è per calcolare la distanza di un punto da un piano:

Definizione 9.3.4. *Dati un punto P e un piano π nello spazio, si definisce la distanza di P da π, $d(P\pi)$, come la distanza di P dal punto d'intersezione P' di π con la retta r, passante per P e perpendicolare a π.*

Se vogliamo calcolare tale distanza, date le coordinate di P e l'equazione cartesiana di π, dovremo trovare la retta r; basterà porre come suo vettore direttore il vettore (a, b, c) perpendicolare al piano, e imporre che passi per P per averne le equazioni parametriche. Sostituendo x, y, z nell'equazione del piano, come nella (9.2), si ottiene un'equazione in t, la cui soluzione determina P', infine la distanza $d(PP')$ ci darà la quantità cercata. Svolgendo il procedimento ora descritto (farlo per esercizio) si ottiene:

Proposizione 9.3.5. *Dati nello spazio un punto $P = (x_P, y_P, z_P)$ e un piano π di equazione $ax + by + cz + d = 0$, si ha che*

$$d(P\pi) = \frac{|ax_P + by_P + cz_P + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

In modo abbastanza analogo si può trattare la distanza fra due piani paralleli.

Definizione 9.3.5. *Dati due piani paralleli π e π', si definisce la distanza di π da π', $d(\pi, \pi')$, come la distanza fra i due punti P, P' intersezione, rispettivamente, di π e di π' con una qualsiasi retta r, perpendicolare ai due piani.*

Se vogliamo calcolare tale distanza, date le equazioni cartesiane dei piani, dovremo trovare una retta r a essi perpendicolare; per far ciò basterà porre come suo vettore direttore il vettore (a, b, c) perpendicolare ai piani per averne le equazioni parametriche (potremo, per semplicità, considerare r passante per l'origine). Sostituendo x, y, z che appaiono in r nell'equazione dei piani, come nella (9.2) si ottengono i due punti P, P', infine la distanza $d(PP')$ ci darà la quantità cercata. Svolgendo il procedimento ora descritto (farlo per esercizio) si ottiene:

Proposizione 9.3.6. *Dati nello spazio due piani paralleli π, π' di equazioni, rispettivamente $ax + by + cz + d = 0$ e $ax + by + cz + d' = 0$, si ha che*

$$d(\pi, \pi') = \frac{|d - d'|}{\sqrt{a^2 + b^2 + c^2}}.$$

Piano parallelo a due rette date

Proposizione 9.3.7. *Siano date nello spazio due rette r e r', non parallele, di parametri direttori rispettivamente l, m, n e l', m', n', e un punto arbitrario $P_0 = (x_0, y_0, z_0)$; allora esiste uno e un solo piano π passante per P_0 e parallelo ad r e r', la cui equazione cartesiana è data da:*

$$\det \begin{pmatrix} x - x_0 & y - y_0 & z - z_0 \\ l & m & n \\ l' & m' & n' \end{pmatrix} = 0.$$

Dimostrazione. Denotiamo con r_0, r'_0 rispettivamente, le rette parallele ad r, r' e passanti per l'origine e consideriamo i punti $R = (l, m, n) \in r_0$, $R' = (l', m', n') \in r'_0$. Osserviamo che, essendo per ipotesi le rette r, r' non parallele, i punti $O = (0, 0, 0)$, $R = (l, m, n)$, $R' = (l', m', n')$ non sono allineati. Pertanto esiste uno e un solo piano che li contiene, ossia uno e un solo piano, diciamo π_0, contenente le rette r_0, r'_0 (parallele rispettivamente a r, r' e passanti per l'origine); π_0 ha equazione (vedi Corollario 9.2.1):

$$\det \begin{pmatrix} x & y & z \\ l & m & n \\ l' & m' & n' \end{pmatrix} = 0.$$

È chiaro che π_0 è l'unico piano passante per l'origine e parallelo a r, r'. Allora il piano π che passa per P_0 ed è parallelo a π_0, conterrà le rette e

sarà il piano cercato. Poiché il piano di equazione:

$$\det \begin{pmatrix} x-x_0 & y-y_0 & z-z_0 \\ l & m & n \\ l' & m' & n' \end{pmatrix} = 0$$

soddisfa tali requisiti (passa per P_0 ed è parallelo a π_0), la proposizione è dimostrata. □

Corollario 9.3.1. *Date nello spazio due rette sghembe r e r', rispettivamente di parametri direttori l, m, n e l', m', n', esistono e sono univocamente determinati due piani paralleli π, π' tali che:*

1. *π contenga r e sia parallelo ad r';*
2. *π' contenga r' e sia parallelo ad r.*

Dimostrazione. Possiamo osservare innanzitutto che se un piano è parallelo a una retta e contiene un punto di essa allora contiene tutta la retta. Per determinare π basterà quindi applicare la Proposizione 9.3.7 considerando un punto $P_0 = (x_0, y_0, z_0)$ appartenente alla retta r, mentre per determinare π' considereremo un punto $P_0' = (x_0', y_0', z_0')$ su r'. □

Fasci di piani

Definizione 9.3.6.

1. Data una retta r dello spazio, si chiama *fascio* (o anche *fascio proprio*) di piani di asse r, l'insieme di tutti i piani dello spazio che contengono r.
2. Dato un piano π, si chiama *fascio improprio* di piani individuato da π, l'insieme di tutti i piani dello spazio paralleli a π.

Proposizione 9.3.8. *Siano dati nello spazio tre piani distinti π, π', π'', di equazione, rispettivamente:*

$\pi : ax+by+cz+d=0$, $\pi' : a'x+b'y+c'z+d'=0$, $\pi'' : a''x+b''y+c''z+d''=0$.

Allora si ha che π, π', π'' appartengono ad uno stesso fascio (proprio o improprio) se e solo se

$$r \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} \leq 2.$$

Dimostrazione. Si ha che i tre piani passano per una retta comune se e solo se il sistema dato dalle loro tre equazioni ha ∞^1 soluzioni, e in tal caso si avrà:

$$r\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = r\begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = 2.$$

Si ha invece che i tre piani sono paralleli se e solo se:

$$r\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = 1.$$

In tal caso avremo sicuramente che il rango della matrice completa è minore o uguale a 2. Viceversa, se il rango della matrice completa è minore o uguale a 2, necessariamente siamo in uno dei due casi descritti sopra. □

Osservazione 9.3.4. Si può introdurre la nozione generale di fascio di piani individuato da due piani distinti π, π', di equazione, rispettivamente:

$$\pi : ax + by + cz + d = 0,\ \pi' : a'x + b'y + c'z + d' = 0$$

definito come l'insieme di tutti i piani dello spazio la cui equazione cartesiana può essere espressa come una combinazione lineare delle equazioni di π, π':

$$l(ax + by + cz + d) + m(a'x + b'y + c'z + d') = 0.$$

Si può notare che quando i due piani non sono paralleli, si ottiene il fascio (proprio) di piani che contengono la retta $\pi \cap \pi'$; se invece i due piani sono paralleli, allora $(a', b', c') = k(a, b, c)$, quindi si ottiene nell'equazione dei piani del fascio:

$l(ax + by + cz + d) + m(kax + kby + kcz + d') =$
$= (l + mk)ax + (l + mk)by + (l + mk)cz + ld + d' = 0,$
e quindi piani paralleli a π (e a π').

Esercizio 9.3.2. Data la retta di equazioni cartesiane

$$r = \begin{cases} 3x - y + z = 0 \\ x - 2y - z + 1 = 0 \end{cases},$$

trovare il piano contenente r e passante per il punto $P = (-1, -1, 17)$.

I piani contenenti r sono dati dal fascio

$$l(3x - y + z) + m(x - 2y - z + 1) = 0.$$

Imponendo il passaggio per P si ottiene:

$$l(-3+1+17)+m(-1+2-17+1)=0,$$

e quindi $15l-15m=0$, e $l=m$. Sostituendo nell'equazione del fascio:

$$m(3x-y+z)+m(x-2y-z+1)=0,$$

e perciò (ponendo $m=1$): $(3x-y+z)+(x-2y-z+1)=0$, da cui $4x-3y+1=0$, che sarà il piano cercato.

9.4 Il prodotto vettoriale

Introduciamo adesso una nuova operazione in $\mathcal{V}_O^3$ che può essere utile in vari problemi geometrici.

Si tratta di un'operazione da $\mathcal{V}_O^3 \times \mathcal{V}_O^3$ in $\mathcal{V}_O^3$ che chiameremo *prodotto vettoriale*.

Definizione 9.4.1. Dati due vettori $\vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathcal{V}_O^3$, si dice *prodotto vettoriale di* $\vec{\mathbf{v}}$ *per* $\vec{\mathbf{w}}$ e si indica con $\vec{\mathbf{v}} \wedge \vec{\mathbf{w}}$ il vettore di $\mathcal{V}_O^3$ così definito:

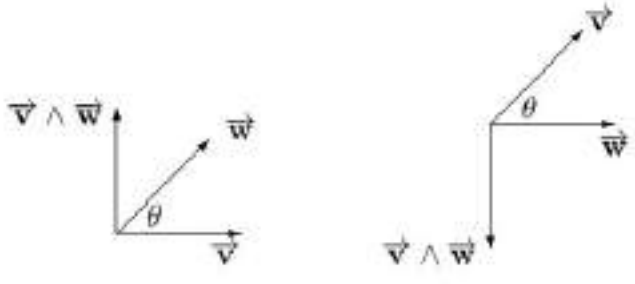
- il modulo di $\vec{\mathbf{v}} \wedge \vec{\mathbf{w}}$ è:

$$|\vec{\mathbf{v}} \wedge \vec{\mathbf{w}}| = |\vec{\mathbf{v}}| \cdot |\vec{\mathbf{w}}| \sin\theta$$

  (ove θ è l'angolo convesso fra $\vec{\mathbf{v}}, \vec{\mathbf{w}}$);

- la direzione di $\vec{\mathbf{v}} \wedge \vec{\mathbf{w}}$ è data dalla retta perpendicolare a $\vec{\mathbf{v}}$ e $\vec{\mathbf{w}}$;

- il verso di $\vec{\mathbf{v}} \wedge \vec{\mathbf{w}}$ è dato dalla "regola della mano destra": ponendo la mano destra di taglio su $\vec{\mathbf{v}}$ in modo che piegando le dita esse vadano in direzione di $\vec{\mathbf{w}}$, il pollice indicherà il verso di $\vec{\mathbf{v}} \wedge \vec{\mathbf{w}}$.



Notiamo che la definizione del modulo di $\vec{v} \wedge \vec{w}$ è tale che si abbia $\vec{v} \wedge \vec{w} = \vec{0}$ se i due vettori sono paralleli (cioè se $\theta = \pi$ oppure $\theta = 0$), oppure se uno di essi è nullo.

Nella definizione della direzione di $\vec{v} \wedge \vec{w}$ potrebbe apparire che essa sia manchevole, in quanto nel caso in cui $\vec{v}$ e $\vec{w}$ siano paralleli non è determinata un'unica retta che sia perpendicolare ad entrambi, ma il problema non si pone in quanto se $\vec{v}$ e $\vec{w}$ sono paralleli, si ha che $\vec{v} \wedge \vec{w} = \vec{0}$, e quindi non è necessario definirne né la direzione né il verso.

Nella seguente proposizione si elencano le proprietà fondamentali del prodotto vettoriale, che si possono dedurre dalla sua definizione.

Proposizione 9.4.1. *Dati i vettori* $\vec{v}$, $\vec{w}$, $\vec{v_1}$, $\vec{v_2}$ *e* $\vec{v_3}$ *in* $\mathcal{V}_O^3$, *valgono le seguenti proprietà:*

1. $\vec{v} \wedge \vec{w} = -\vec{w} \wedge \vec{v}$;
2. $\vec{v_1} \wedge (\vec{v_2} + \vec{v_3}) = \vec{v_1} \wedge \vec{v_2} + \vec{v_1} \wedge \vec{v_3}$;
3. $(\lambda \vec{v_1}) \wedge \vec{v_2} = \lambda (\vec{v_1} \wedge \vec{v_2}) = \vec{v_1} \wedge (\lambda \vec{v_2})$, $\forall\, \lambda \in \mathbb{R}$.

L'utilità del prodotto vettoriale è in primo luogo il fatto che dati due vettori ne fornisce un terzo ortogonale ad entrambi; tuttavia perché esso sia utilizzabile in problemi di geometria analitica dello spazio abbiamo bisogno di averne una "versione in $\mathbb{R}^3$", in quanto vogliamo sfruttare il metodo delle coordinate (come abbiamo fatto per il prodotto scalare) per poter agevolmente "fare i conti" con i vettori traducendoli come elementi di $\mathbb{R}^3$.

Consideriamo come sempre un sistema di riferimento cartesiano ortogonale $Oxyz$ in $\mathcal{V}_O^3$; cioè fissiamo in V_O^3 tre vettori ortogonali $\left(\vec{i}, \vec{j}, \vec{k}\right)$ di modulo 1, come in Figura 9.7:

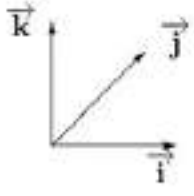


Figura 9.7: Sistema di riferimento cartesiano ortogonale

Notiamo che si ha

$$\vec{i} \wedge \vec{j} = \vec{k}, \ \vec{i} \wedge \vec{k} = -\vec{j} \ \vec{j} \wedge \vec{k} = \vec{i}.$$

Osservazione 9.4.1. Nella scelta della base $\left(\vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}}\right)$, conta anche che essi siano posizionati come in figura, perché se ad esempio invertissimo le posizioni di due di essi, anche i segni dei prodotti vettoriali cambierebbero.

Ogni vettore $\overrightarrow{\mathbf{OP}} \in \mathcal{V}_O^3$ si scriverà in modo unico nella forma $\overrightarrow{\mathbf{OP}} = x\vec{\mathbf{i}} + y\vec{\mathbf{j}} + z\vec{\mathbf{k}}$; allora se abbiamo due vettori $\vec{\mathbf{v}} = \overrightarrow{\mathbf{OP}} = \left(a\vec{\mathbf{i}} + b\vec{\mathbf{j}} + c\vec{\mathbf{k}}\right)$ e $\vec{\mathbf{w}} = \overrightarrow{\mathbf{OQ}} = \left(a'\vec{\mathbf{i}} + b'\vec{\mathbf{j}} + c'\vec{\mathbf{k}}\right)$ varrà:

$$\vec{\mathbf{v}} \wedge \vec{\mathbf{w}} = \overrightarrow{\mathbf{OP}} \wedge \overrightarrow{\mathbf{OQ}} = \left(a\vec{\mathbf{i}} + b\vec{\mathbf{j}} + c\vec{\mathbf{k}}\right) \wedge \left(a'\vec{\mathbf{i}} + b'\vec{\mathbf{j}} + c'\vec{\mathbf{k}}\right) =$$

$$= ab'\vec{\mathbf{k}} - ac'\vec{\mathbf{j}} - ba'\vec{\mathbf{k}} + bc'\vec{\mathbf{i}} + ca'\vec{\mathbf{j}} - cb'\vec{\mathbf{i}} =$$

$$= (bc' - cb')\vec{\mathbf{i}} + (ca' - ac')\vec{\mathbf{j}} + (ab' + ba')\vec{\mathbf{k}}.$$

Possiamo adesso "tradurre" l'operazione di prodotto vettoriale anche come una operazione in $\mathbb{R}^3$, ponendo:

$$\forall\, (a,b,c), (a',b',c') \in \mathbb{R}^3 : \ (a,b,c) \wedge (a',b',c') = (bc'-cb', ca'-ac', ab'-ba').$$

Esempio 9.4.1. $(2,1,0) \wedge (1,-1,1) = (1,-2,-3)$,
$(1,0,0) \wedge (0,0,2) = (0,-2,0)$.

Osserviamo che un modo per memorizzare la formula del prodotto vettoriale è quello di esprimerla in forma determinantale:

$$(a,b,c) \wedge (a',b',c') = (bc'-cb', ca'-ac', ab'-ba') = \det \begin{pmatrix} \vec{\mathbf{i}} & \vec{\mathbf{j}} & \vec{\mathbf{k}} \\ a & b & c \\ a' & b' & c' \end{pmatrix}.$$

Notare che il determinante scritto qui sopra è "finto", in quanto non si tratta di una vera e propria matrice (la prima riga ha per elementi dei vettori non dei numeri); però la formula del determinante applicata a questa "finta matrice" dà un utile modo di ricordare la formula del prodotto vettoriale.

Applicazioni

- Se abbiamo una retta in forma cartesiana:

$$r : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

e vogliamo un suo vettore direzionale, consideriamo dapprima i due vettori (a,b,c) e (a',b',c'), essi sono perpendicolari a due piani che definiscono la retta, e quindi sono anche perpendicolari a r; si ha allora che un vettore perpendicolare a essi sarà parallelo alla retta

(quindi utilizzabile come suo vettore direzionale). Avremo così che il vettore:

$$(a,b,c) \wedge (a',b',c') = (bc' - cb', ca' - ac', ab' - ba')$$

è un vettore direzionale per r.

- Se abbiamo due rette non parallele:

$$\begin{cases} x = lt + x_0 \\ y = mt + y_0 \\ z = nt + z_0 \end{cases} , \quad \begin{cases} x = l't + x_1 \\ y = m't + y_1 \\ z = n't + z_1 \end{cases}$$

e vogliamo un piano $ax + by + cz + d = 0$, parallelo a entrambe, dovremo avere che il vettore (a,b,c) sia perpendicolare a entrambe le rette; potremo prendere allora:

$$(a,b,c) = (l,m,n) \wedge (l',m',n')$$

e avremo un piano come desiderato (naturalmente d può essere qualsiasi).

9.5 Esercizi

In tutti gli esercizi che seguono l'ambiente in cui si lavora è lo spazio tridimensionale ($\mathbb{R}^3$ o $\mathcal{V}_O^3$).

Esercizio 9.5.1. Scrivere l'equazione del piano passante per i punti $A : (0,1,0)$, $B : (2,1,0)$ e $C : (0,1,1)$.

Esercizio 9.5.2. Scrivere l'equazione del piano π passante per il punto $A : (2,1,-1)$ e parallelo al piano π' definito dall'equazione cartesiana $x - 3y + z - 1 = 0$.

Esercizio 9.5.3. Disegnare il piano dell'esercizio precedente.

Esercizio 9.5.4. Scrivere l'equazione del piano di $\mathbb{R}^3$ contenente la retta $x = y = z - 1$ e il punto $P : (5,1,1)$.

Esercizio 9.5.5. Scrivere l'equazione del piano π passante per i punti $A : (3,0,1)$, $B : (2,2,2)$ e parallelo alla direzione individuata dal vettore $\overline{\mathbf{v}} = (1,1,0)$.

Esercizio 9.5.6. Scrivere l'equazione del piano π passante per i punti $P : (1,1,1)$, $Q : (2,1,2)$ e parallelo alla retta r di equazioni

$$\begin{cases} x = t \\ y = 2t \\ z = 3t \end{cases} .$$

Esercizio 9.5.7. Determinare la distanza tra i piani π e π', di equazioni rispettive $x - y + 4z = 0$ e $x - y + 4z - 9 = 0$.

Esercizio 9.5.8. Disegnare i piani dell'esercizio precedente.

Esercizio 9.5.9. Scrivere l'equazione del piano assiale al segmento di estremi $A : (1,0,1)$ e $B : (2,2,0)$ (Il piano assiale passa per il punto medio del segmento AB ed è a esso perpendicolare).

Esercizio 9.5.10. Si considerino, nello spazio, i punti:
$A : (1,1,1)$, $B : (0,1,0)$, $C : (0,2,1)$, $D : (0,0,0)$.
Stabilire se siano complanari.

Esercizio 9.5.11. Siano date le rette

$$r : \begin{cases} x + y + z + 1 = 0 \\ x - y = 1 \end{cases}, \quad s : \begin{cases} x = 1 + 2t \\ y = 2 + 3t \\ z = t \end{cases}.$$

Scrivere le equazioni parametriche di r e le equazioni cartesiane di s.

Esercizio 9.5.12. Scrivere le equazioni della retta congiungente i punti $A : (1,1,1)$ e $B : (2,3,5)$.

Esercizio 9.5.13. Scrivere le equazioni di una retta r che passi per il punto $Q : (1,-1,0)$, e sia parallela al piano $\pi : \ 2x - y + 4 = 0$.

Esercizio 9.5.14. Scrivere l'equazione del piano passante per il punto $P : (1,1,0) \in \mathbb{R}^3$ e parallelo al piano di equazione $x - z + 1 = 0$.

Esercizio 9.5.15. Determinare la retta r passante per i punti $A : (1,2,3)$ e $B : (0,-4,1)$.

Esercizio 9.5.16. Scrivere le equazioni della retta r passante per il punto $P : (1,0,0)$ e ortogonale al piano π descritto in forma parametrica dalle seguenti equazioni:

$$\pi : \begin{cases} x = 2 + t_1 \\ y = t_2 \\ z = 3 + t_1 \end{cases}.$$

Esercizio 9.5.17. Sia r la retta così definita:

$$r : \begin{cases} x = 3 + t \\ y = -t \\ z = -2 - 2t \end{cases};$$

trovare una retta passante per il punto $B : (3,1,-2)$, sghemba con r e parallela al piano π:

$$\pi : \begin{cases} x = 1 - t_1 + t_2 \\ y = 3t_2 \\ z = -t_1 - 2t_2 \end{cases}.$$

Esercizio 9.5.18. Date le rette r, $r' \subset \mathbb{R}^3$ così definite:

$$r: \begin{cases} x = 1 + y \\ z = 0 \end{cases} \quad \text{e } r': \begin{cases} x = 1 + kz \\ y = 0 \end{cases},$$

determinare la loro posizione reciproca (sghembe, parallele, incidenti) al variare di k.

Esercizio 9.5.19. Scrivere l'equazione del piano passante per il punto $P : (1, 2, 0)$ e contenente la retta r definita dalle seguenti equazioni cartesiane:

$$r: \begin{cases} x + y = 0 \\ 2x + z + 1 = 0 \end{cases}.$$

Esercizio 9.5.20. Scrivere l'equazione della retta r passante per il punto $Q : (1/2, -3, 0)$, parallela al piano di equazione cartesiana $2x - y + 1 = 0$ e incidente alla retta s:

$$s: \begin{cases} x = y \\ y = z + 2 \end{cases}.$$

Esercizio 9.5.21. Stabilire se le rette $r: \begin{cases} y = 1 \\ z = 3 \end{cases}$ e $s: \begin{cases} x = 1 \\ y - z + 1 = 0 \end{cases}$ siano incidenti, parallele o sghembe.

Esercizio 9.5.22. Disegnare le rette dell'esercizio precedente.

Esercizio 9.5.23. Stabilire per quali valori del parametro reale h le rette

$$r: \begin{cases} x - 2hy + 3z = 0 \\ hx + y = 0 \end{cases}, \quad s: \begin{cases} x + hy = 2 \\ y - hz = 0 \end{cases}$$

sono parallele e determinare se siano disgiunte o coincidenti.

Esercizio 9.5.24. Scrivere le equazioni della retta r che passa per il punto $P : (1, 1, 1)$, è perpendicolare alla retta s di equazioni:

$$\begin{cases} x = 2 + t \\ y = 2 + 2t \\ z = 2 + 3t \end{cases}$$

ed è parallela al piano π di equazione $x + 3y = 0$.

Esercizio 9.5.25. Scrivere le equazioni della retta incidente e perpendicolare ad entrambe le rette

$$r: \begin{cases} x = 1 \\ y = 0 \\ z = \tau \end{cases}, \quad s: \begin{cases} x = t \\ y = 2t \\ z = 2t \end{cases}.$$

Calcolare la distanza minima tra r e s.

Esercizio 9.5.26. Stabilire se le seguenti rette sono complanari o sghembe:

$$r: \begin{cases} x-1=0 \\ y=0 \end{cases}, \quad s: \begin{cases} 2x=y \\ y=z \end{cases}.$$

Esercizio 9.5.27. Siano date nello spazio le rette sghembe

$$r: \begin{cases} 2x-z=0 \\ y-z=0 \end{cases}, \quad s: \begin{cases} x+y+z-1=0 \\ 3x+2z=0 \end{cases}.$$

Determinare un piano parallelo ad entrambe.

Esercizio 9.5.28. Calcolare la distanza dal punto $P:(2,0,1)$ dal piano π di equazione:

$$3x-7y-5z-1=0.$$

Esercizio 9.5.29. Scrivere l'equazione del luogo dei punti dello spazio equidistanti dal piano $\pi: z=1$ e dal punto $P:(2,2,0)$.

Esercizio 9.5.30. Siano dati i piani π, π' di equazioni rispettive:
$2x+y+z=0$ e $x-z-1=0$.
Scrivere l'equazione del piano π'' simmetrico di π' rispetto a π.

Esercizio 9.5.31. Si consideri il sistema

$$\begin{cases} (h+1)x+y-h-2=0 \\ y+hz-1=0 \end{cases}$$

dove $h \in \mathbb{R}$.

1. Verificare che $\forall\ h \in \mathbb{R}$, il sistema rappresenta una retta nello spazio;
2. detta $\mathcal{F}_h$ la famiglia di rette data dal sistema al variare di h, determinare una retta di $\mathcal{F}_h$ perpendicolare alla retta
$$r: 3x=-3y=-2z.$$

Esercizio 9.5.32. Si consideri la famiglia di rette $\mathcal{F}_h$ dell'esercizio precedente

1. stabilire se esistono rette di $\mathcal{F}_h$ ortogonali al piano π di equazione
$$x-2y+2z-3=0;$$
2. verificare che le tutte rette di $\mathcal{F}_h$ passano per uno stesso punto V. Determinare V.

Esercizio 9.5.33. Si considerino la famiglia di piani

$$\mathcal{F}: (k^2-2k+2)x+(2-k)y+z-1=0$$

al variare di k, e la retta $r: 2x=2y=-z$.

1. Determinare i piani di $\mathcal{F}$ paralleli a r.

2. Determinare i piani di $\mathcal{F}$ perpendicolari a r.

Esercizio 9.5.34. Si considerino le rette

$$r: \begin{cases} x = 1 - t \\ y = h + t \\ z = t \end{cases}, \quad s: \begin{cases} x + y - hz = 0 \\ 2x + hz - 1 = 0 \end{cases}.$$

Trovare i valori di h in modo che r e s siano incidenti. In corrispondenza dei valori trovati scrivere l'equazione del piano contenente r e s.

Esercizio 9.5.35. Dati nello spazio il piano $kx + 2y - (3+k)z + 41 = 0$ e la retta

$$r: \begin{cases} x = 3t \\ y = t \\ z = 2t \end{cases};$$

determinare la posizione reciproca di piano e retta al variare di k e per quali valori di k, se esistono, si abbia che piano e retta siano tra loro perpendicolari.

Esercizio 9.5.36. Trovare, se possibile, $k \in \mathbb{R}$ in modo tale che i seguenti punti siano complanari: $P_1 : (0, 0, k)$, $P_2 : (1, k, 0)$, $P_3 : (k, 0, k)$ e $P_4 : (1, 1, k)$. Dopodiché, per uno di questi valori di k, trovare un piano parallelo a quello generato da $P_1, \ldots, P_4$ e passante per $Q = (1, 1, 1)$.

Esercizio 9.5.37. Sia dato il fascio di piani incidenti in

$$r: \begin{cases} x = 2 + t \\ y = -t \\ z = 1 - t \end{cases}.$$

Trovare in esso il piano π passante per il punto $P : (0, 0, 2)$. Si determini poi il piano π' parallelo a π e passante per $Q : (2, 4, 0)$ e la retta r passante per Q e perpendicolare a π.

Esercizio 9.5.38. * Fare un esempio di due rette r, r' che siano incidenti e uno in cui r, r' siano parallele disgiunte. In entrambi i casi, determinare il piano che contiene le due rette. [Le rette possono essere date in forma parametrica o cartesiana, come si preferisce].

Esercizio 9.5.39. * Fare un esempio di una rette r e un piano π che siano paralleli e disgiunti. Determinare poi la distanza fra retta e piano.

Esercizio 9.5.40. * Considerare la retta

$$r: \begin{cases} x = t + 1 \\ y = 2t + 2 \\ z = t - 3 \end{cases}.$$

Lanciando tre dadi, sia $(a, b, c) \in \{1, 2, 3, 4, 5, 6\}^3$ il risultato ottenuto, e consideriamo la retta:

$$r': \begin{cases} x = as \\ y = bs \\ z = cs \end{cases}.$$

Dire se sia più probabile ottenere che:

- r e r' siano sghembe;
- r e r' siano incidenti;
- r e r' siano parallele disgiunte;
- $r = r'$.

CAPITOLO 10

Forme bilineari e quadratiche

Prerequisiti: Fondamenti della geometria analitica nel piano e nello spazio; rango di una matrice, determinante; isometrie; calcolo degli autovalori e degli autovettori; teorema spettrale.

Obiettivi: Saper riconoscere, utilizzare e classificare forme bilineari e forme quadratiche; applicare lo studio di nozioni di algebra lineare apprese nei capitoli precedenti al problema geometrico dello studio e della classificazione delle curve piane dette coniche e delle superfici nello spazio dette quadriche (entrambe definite tramite equazioni polinomiali di secondo grado).

10.1 Forme bilineari

Introduciamo adesso un concetto, quello di forma bilineare, che ci consentirà di generalizzare la nozione di prodotto scalare.

Definizione 10.1.1. Una *forma bilineare* su uno spazio vettoriale reale V è una funzione:

$$\varphi : V \times V \to \mathbb{R}$$

tale che per ogni $v, v', w, w' \in V$ e per ogni $\alpha \in \mathbb{R}$

1. $\varphi(v + v', w) = \varphi(v, w) + \varphi(v', w)$;
2. $\varphi(v, w + w') = \varphi(v, w) + \varphi(v, w')$;
3. $\varphi(\alpha v, w) = \varphi(v, \alpha w) = \alpha\varphi(v, w)$.

Si può osservare che per ogni forma bilineare si ha, $\forall v \in V$:

$$\varphi(0_V, v) = \varphi(v, 0_V) = 0.$$

Esempio 10.1.1. Sia $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definita da

$$\varphi((x_1, x_2), (y_1, y_2)) = (x_1 y_2 + 3x_2(y_1 + y_2)).$$

Essa è una forma bilineare. Infatti se $v = (x_1, x_2)$, $v' = (x_1', x_2')$ e $w = (y_1, y_2)$ abbiamo che $\varphi(v + v', w) = \varphi((x_1 + x_1', x_2 + x_2'), (y_1, y_2)) =$
$$= ((x_1 + x_1')y_2 + 3(x_2 + x_2')(y_1 + y_2)) =$$
$= ((x_1 y_2 + 3x_2(y_1 + y_2)) + (x_1' y_2 + 3x_2'(y_1 + y_2))) = \varphi(v, w) + \varphi(v', w)$ che verifica la prima proprietà. Le altre due proprietà si verificano in modo del tutto analogo e le lasciamo per esercizio.

Se V ha dimensione finita n, allora ogni forma bilineare φ su V può essere rappresentata come una matrice quadrata con n righe. Così come si faceva per le forme lineari, occorre fisare una base $B = \{v_1, \ldots, v_n\}$ di V (la matrice associata dipenderà infatti dalla base scelta).

La matrice $M_B(\varphi) = (b_{i,j})_{i,j=1,\ldots,n}$ è così definita:

$$b_{i,j} = \varphi(v_i, v_j), \quad \text{per ogni } i, j = 1, \ldots, n.$$

Quando sia nota la matrice associata ad una certa forma bilineare φ (e la base rispetto a cui è rappresentata), l'azione di φ su due vettori u e w di V si può quindi ricavare nel modo seguente:

$$\varphi(u, w) = u^T M_B(\varphi) w = \sum_{i,j=1}^{n} b_{ij} u_i w_j \tag{10.1}$$

dove u_i e w_j, $i, j = 1, \ldots, n$, sono le coordinate di u e w rispetto alla base B.

Esempio 10.1.2. La matrice associata alla forma bilineare dell'esempio precedente $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$,

$$\varphi((x_1, x_2), (y_1, y_2)) = (x_1 y_2 + 3x_2(y_1 + y_2)),$$

rispetto alla base canonica $\mathcal{E}$, è

$$M_{\mathcal{E}}(\varphi) = \begin{pmatrix} 0 & 1 \\ 3 & 3 \end{pmatrix}.$$

Teorema 10.1.1. *Due matrici quadrate $M_B(\varphi)$ e $M_{B'}(\varphi)$ rappresentano la stessa forma bilineare (rispetto a due basi B e B' eventualmente distinte) se e solo se*

$$M_B(\varphi) = (M_{B,B'}(id))^T M_{B'}(\varphi) M_{B,B'}(id),$$

dove $M_{B,B'}(id)$ e $M_{B,B'}(id)$ sono le matrici di cambiamento di base rispettivamente da B a B' e viceversa.

Dimostrazione. Infatti
$\varphi(v'_B, w'_B) = v'^T_B M_{B'}(\varphi) w'_B =$
$= (M_{B,B'}(id)(v_B))^T M_{B'}(\varphi)(M_{B,B'}(id)w_B) =$
$v_B^T (M_{B,B'}(id))^T M_{B'}(\varphi) M_{B,B'}(id) w_B = \varphi(v_B, w_B).$
Quindi $M_B(\varphi) = (M_{B,B'}(id))^T M_{B'}(\varphi) M_{B,B'}(id)$. □

Definizione 10.1.2. *Una forma bilineare simmetrica* su V è una forma bilineare $\varphi : V \times V \to \mathbb{R}$ tale che

$$\varphi(v, w) = \varphi(w, v)$$

per ogni $v, w \in V$.

Esempio 10.1.3. La seguente forma bilineare $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definita da

$$\varphi((x_1, x_2), (y_1, y_2)) = (x_1 y_1 - 2x_2 y_1 - 2x_1 y_2)$$

è simmetrica infatti $\varphi((x_1, x_2), (y_1, y_2)) = (x_1 y_1 - 2x_2 y_1 - 2x_1 y_2)$ e $\varphi((y_1, y_2), (x_1, x_2)) = (y_1 x_1 - 2y_2 x_1 - 2y_1 x_2)$.

Esempio 10.1.4. La forma bilineare degli Esempi 10.1.1 e 10.1.2 non è simmetrica.

Osservazione 10.1.1. Una forma φ è simmetrica se e solo se la matrice associata $M_B(\varphi)$ rispetto a una qualsiasi base B è simmetrica.

Esempio 10.1.5. La matrice associata alla forma bilineare dell'esempio precedente $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definita da

$$\varphi((x_1, x_2), (y_1, y_2)) = (x_1 y_1 - 2x_2 y_1 - 2x_1 y_2)$$

rispetto alla base canonica $\mathcal{E}$ è

$$M_{\mathcal{E}}(\varphi) = \begin{pmatrix} 1 & -2 \\ -2 & 0 \end{pmatrix}.$$

L'espressione matriciale (10.1) di una forma bilineare simmetrica può quindi essere riscritta:

$$\varphi(u, w) = u^T M_B(\varphi) w = w^T M_B(\varphi) u.$$

Dal Teorema 10.1.1 segue quindi il seguente

Corollario 10.1.1. *Una matrice associata a una forma bilineare simmetrica è sempre una matrice simmetrica indipendentemente dalla scelta della base.*

Definizione 10.1.3. Sia φ una forma bilineare simmetrica su V. Due vettori $v, w \in V$ si dicono *ortogonali* rispetto a φ se $\varphi(v,w) = 0$. Un vettore $v \in V$ si dice *isotropo* per φ se $\varphi(v,v) = 0$.

Esempio 10.1.6. Un esempio di vettore isotropo per la forma bilineare degli Esempi 10.1.3 e 10.1.5 $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definita da

$$\varphi((x_1,x_2),(y_1,y_2)) = (x_1y_1 - 2x_2y_1 - 2x_1y_2)$$

è $v = (0,1)$. Infatti $\varphi(v,v) = \varphi((\mathbf{0},\mathbf{1}),(\mathbf{0},\mathbf{1})) = (\mathbf{0}\cdot\mathbf{1} - \mathbf{2}\cdot\mathbf{1}\cdot\mathbf{0} - \mathbf{2}\cdot\mathbf{0}\cdot\mathbf{1}) = (\mathbf{0})$.

Definizione 10.1.4. Per ogni sottospazio $U \subset V$ si può definire il suo *sottospazio ortogonale* rispetto a una forma bilineare simmetrica φ:

$$U^{\perp_\varphi} := \{v \in V \mid \varphi(u,v) = 0, \forall\, u \in U\}.$$

Esempio 10.1.7. Consideriamo la seguente forma bilineare simmetrica $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definita da

$$\varphi((x_1,x_2),(y_1,y_2)) = (x_1y_2 + x_2y_1).$$

E cerchiamo l'ortogonale rispetto a φ di $U = \langle (1,0)_{\mathcal{E}} \rangle$.

$U^{\perp_\varphi} = \{v = (z_1,z_2) \in \mathbb{R}^2 \mid \varphi((1,0),(z_1,z_2)) = 0\}$.

Ora $\varphi((1,0),(z_1,z_2)) = z_2$, quindi $U^{\perp_\varphi} = \{v = (z_1,0) \in \mathbb{R}^2\} = \langle (1,0) \rangle$. Questo è un caso particolare in cui accade che $U = U^{\perp_\varphi}$, questo fenomeno ovviamente non accade sempre, anzi è piuttosto speciale.

Si possono semplicemente verificare le seguenti proprietà.

Siano $U_1, U_2 \subset V$ due sottospazi di V e φ una forma bilineare simmetrica su V, allora

1. se $U_1 \subseteq U_2$, allora $U_2^{\perp_\varphi} \subseteq U_1^{\perp_\varphi}$;
2. $(U_1 + U_2)^{\perp_\varphi} = U_1^{\perp_\varphi} \cap U_2^{\perp_\varphi}$;
3. $(U_1 \cap U_2)^{\perp_\varphi} = U_1^{\perp_\varphi} + U_2^{\perp_\varphi}$.

Definizione 10.1.5. Si dice *radicale* di φ, forma bilineare simmetrica su V, il seguente insieme:

$$V^{\perp_\varphi} := \{v \in V \mid \varphi(v,w) = 0, \forall\, w \in V\}.$$

Esempio 10.1.8. Consideriamo la seguente forma bilineare simmetrica $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathrm{R}$ definita da

$$\varphi((x_1, x_2), (y_1, y_2)) = (x_1 y_1)$$

e cerchiamo il suo radicale.
$V^{\perp_\varphi} = \{v = (x_1, x_2) \in \mathbb{R}^2 \mid \varphi((x_1, x_2), (y_1, y_2)) = 0, \forall (y_1, y_2) \in \mathbb{R}^2\}$. Ora $\varphi((x_1, x_2), (y_1, y_2)) = (x_1 y_1)$ quindi avremo $x_1 y_1 = 0$ comunque scelga (y_1, y_2) se e solo se prendiamo $x_1 = 0$, perciò $V^{\perp_\varphi} = \langle (0, 1) \rangle$.

Definizione 10.1.6. Una forma bilineare simmetrica si dice *non degenere* se il suo radicale è nullo, *degenere* viceversa.

Una forma bilineare simmetrica φ si dice *definita positiva* se, per ogni $v \in V - \{0_V\}$, si ha $\varphi(v, v) > 0$.

Esempio 10.1.9. È un semplice esercizio mostrare che la forma bilineare dell'Esempio 10.1.7 è non degenere mentre è ovvio da quanto visto nell'esempio precedente che la forma bilineare dell'Esempio 10.1.8 è degenere.

Osservazione 10.1.2. Una forma bilineare simmetrica definita positiva è ovviamente non degenere.

Se ricordiamo la definizione di **prodotto scalare** ci accorgiamo che essa coincide con la definizione di **forma bilineare simmetrica definita positiva**. Quindi questa potrebbe essere anche presa come definizione alternativa di prodotto scalare.

Proposizione 10.1.1. *La forma bilineare simmetrica φ è non-degenere se e solo se la matrice associata a φ rispetto a una qualsiasi base di V è di rango massimo.*

Dimostrazione. Fissiamo un vettore $v \in V$ e consideriamo l'applicazione

$$b(v, \cdot) : V \to \mathbb{R}$$

$$w \mapsto b(v, w).$$

Questa è chiaramente una forma lineare su V, inoltre $v \in V^{\perp_\varphi}$ se e solo se $b(v, \cdot)$ è la forma lineare nulla. La matrice associata all'applicazione lineare $b(v, \cdot)$ rispetto a una base B coincide col vettore riga $v_B^T M_B(\varphi)$, perciò $v \in V^{\perp_\varphi}$ se e solo se tale vettore riga è zero, ossia se e solo se il vettore colonna $M_B(\varphi) v_B$ è zero, ossia il vettore colonna v_B appartiene a un sottospazio di dimensione $\dim(V) - rg(M_B(\varphi))$. □

Esempio 10.1.10. È un semplice esercizio mostrare che ogni matrice associata alla forma bilineare dell'Esempio 10.1.7 è di rango 2 mentre ogni matrice associata alla forma bilineare dell'Esempio 10.1.8 è di rango 1.

Teorema 10.1.2 (Teorema di Sylvester). *Sia V uno spazio vettoriale di dimensione finita n, sia φ una forma bilineare simmetrica su V. Allora esiste una base B di V tale che la matrice associata alla forma bilineare φ rispetto alla base B sia della seguente forma a blocchi*

$$\left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0 \end{array}\right) \tag{10.2}$$

con $p, q \geq 0$, $p+q \leq n$, e p, q NON dipendono dalla base B scelta ma solo da φ.

Dimostrazione. Iniziamo col dimostrare che data una forma bilineare simmetrica su V, è sempre possibile trovare una base B di V rispetto alla quale $M_B(\varphi)$ sia diagonale.

Si vede per induzione. Se $n = 1$ è ovvio.

Supponiamo che sia vero per ogni spazio vettoriale di dimensione $n-1$, e dimostriamolo per lo spazio vettoriale V di dimensione n con φ forma bilineare simmetrica su V. Se $Im(\varphi) = \{0\}$ allora tutte le basi di V vanno bene. Supponiamo che φ non sia identicamente nulla, prendiamo un vettore v_1 non isotropo. Osserviamo che questo è sempre possibile in quanto essendo φ non è identicamente nulla, esistono necessariamente due vettori $u_1, u_2 \in V$ tali che $\varphi(u_1, u_2) \neq 0$. Ora se u_1 e u_2 fossero entrambi isotropi, allora basterebbe prendere $u_1 + u_2$ come vettore non isotropo:

$$\varphi(u_1+u_2, u_1+u_2) = \varphi(u_1,u_1) + 2\varphi(u_1,u_2) + \varphi(u_2,u_2) = 2\varphi(u_1,u_2) \neq 0.$$

Sia quindi $v_1 \in V$ un vettore non isotropo e consideriamo $\{v_1\}^{\perp_\varphi}$. Quest'ultimo spazio ha dimensione $n-1$ in quanto coincide col nucleo della mappa lineare $\varphi(v_1, \cdot)$ introdotta all'inizio della dimostrazione della Proposition 10.1.1. Tale applicazione lineare è non nulla. Inoltre $v_1 \notin \{v_1\}^{\perp_\varphi}$. Se consideriamo la restrizione $\varphi|_{\{v_1\}^{\perp_\varphi}}$ di φ al sottospazio $\{v_1\}^{\perp_\varphi}$ essa risulta essere una forma bilineare simmetrica definita su uno spazio $(n-1)$-dimensionale. Quindi possiamo applicare l'ipotesi induttiva e affermare che esiste una base $B' = v_2, \ldots, v_n$ di $\{v_1\}^{\perp_\varphi}$ rispetto alla quale $M_{B'}\left(\varphi|_{\{v_1\}^{\perp_\varphi}}\right)$ è diagonale. A questo punto è se prendiamo $B = v_1, v_2, \ldots, v_n$ come base di V otteniamo che

$$M_B(\varphi) = \left(\begin{array}{c|c} \varphi(v_1,v_1) & 0 \\ \hline 0 & M_{B'}\left(\varphi|_{\{v_1\}^{\perp_\varphi}}\right) \end{array}\right).$$

Abbiamo così raggiunto il nostro primo obiettivo di dimostrare che esiste una base di V rispetto alla quale la matrice associata a una forma bilinerare simmetrica è diagonale.

Ora chiamiamo $m_i := \varphi(v_i, v_i)$, per $i = 1, \ldots, n$. Essendo m_i numeri reali per definizione di φ, essi possono essere positivi, negativi o nulli. Supponiamo che siano ordinati in modo tale che per qualche $p, q \geq 0$ si abbia che $m_1, \ldots, m_p > 0$, $m_{p+1}, \ldots, m_{p+q} < 0$ e $m_{p+q+1}, \ldots, m_n = 0$ con $p+q \leq n$. Ora è un semplice esercizio verificare che la base $\bar{B} = \left\{\frac{v_1}{\sqrt{m_1}}, \ldots, \frac{v_p}{\sqrt{m_p}}, \frac{v_{p+1}}{\sqrt{-m_{p+1}}}, \ldots, \frac{v_{p+q}}{\sqrt{-m_{p+q}}}, v_{p+q+1}, \ldots, v_n\right\}$

è tale che

$$M_B(\varphi) = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

Resta quindi da dimostrare che p, q non dipendono dalla scelta dalla base. Innanzitutto osserviamo che la somma $p+q$ coincide col rango della matrice che sappiamo già essere indipendente dalla scelta della base. È quindi sufficiente dimostrare che uno dei due non dipende dalla scelta della base. Mostriamo che p non dipende dalla base scelta, perché è la dimensione massima di un sottospazio vettoriale di V su cui la restrizione di φ sia definita positiva. Sia $W \subset V$ un altro sottospazio su cui la restrizione di φ è definita positiva e sia $U = \langle v_{p+1}, \ldots, v_n \rangle$. Se $\dim(W) > p$ avremmo per la formula di Grassmann che

$$\dim(W \cap U) = \dim W + \dim U - \dim(W + U) > p + (n - p) - n = 0$$

ossia $W \cap \langle v_{p+1}, \ldots, v_n \rangle \neq \emptyset$ che è assurdo. □

Esempio 10.1.11. Consideriamo la seguente forma bilineare simmetrica $\varphi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ definita da

$$\varphi((x_1, x_2, x_3), (y_1, y_2, y_3)) = (x_1y_1 + 2x_1y_2 + 2x_2y_1 + 4x_2y_2 - x_3y_3).$$

La matrice associata a φ rispetto alla base canonica è

$$M_{\mathcal{E}}(\varphi) = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Gli autovalori di $M_{\mathcal{E}}(\varphi)$ sono $\lambda_1 = 5, \lambda_2 = -1$ e $\lambda_3 = 0$ i cui rispettivi autovettori sono $v_{\lambda_1} = (1, 2, 0), v_{\lambda_2} = (0, 0, 1)$ e $v_{\lambda_3} = (-2, 1, 0)$. Perciò se scegliamo la base $B = \{v_{\lambda_1}, v_{\lambda_2}, v_{\lambda_3}\}$ abbiamo che

$$M_B(\varphi) = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

La dimostrazione del teorema di Sylvester mostra che basta prendere $B' = \left\{\frac{v_{\lambda_1}}{\sqrt{5}}, v_{\lambda_2}, v_{\lambda_3}\right\}$ per avere

$$M_{B'}(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

che è propio nella forma a blocchi (10.2) con $p = q = 1$.

Definizione 10.1.7. La coppia $(p, q) \in \mathbb{N} \times \mathbb{N}$ del teorema di Sylvester si chiama *segnatura* della forma bilineare simmetrica φ.

Corollario 10.1.2. *Una forma bilineare simmetrica definita su uno spazio vettoriale di dimensione n che abbia segnatura (p, q) è non-degenere se e solo se $p+q=n$, ed è definita positiva se e solo se $p=n$.*

10.2 Forme quadratiche

Definizione 10.2.1. Sia φ una forma bilineare simmetrica su uno spazio vettoriale V. Si definisce Q la *forma quadratica associata a* φ la funzione:

$$Q : V \to \mathbb{R}$$

$$v \mapsto \varphi(v, v).$$

Esempio 10.2.1. Sia $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definita da

$$\varphi((x_1, x_2), (y_1, y_2)) = -x_1 y_2 - x_2 y_1 + x_2 y_2$$

allora la forma quadratica $Q : \mathbb{R}^2 \to \mathbb{R}$ associata è definita come

$$Q((x_1, x_2)) = \varphi((x_1, x_2), (x_1, x_2)) = -2x_1 x_2 + x_2^2.$$

Teorema 10.2.1. *Sia Q una forma quadratica su uno spazio vettoriale V associata a una forma bilineare φ. Allora :*

1. $Q(v+w) = Q(v) + 2\varphi(x, y) + Q(y)$;
2. $Q(\alpha v) = \alpha^2 Q(v)$

per ogni $v, w \in V$ e per ogni $\alpha \in \mathbb{R}$.

Dimostrazione. 1. La prima proprietà è una conseguenza del fatto che $Q(v+w) = \varphi(v+w, v+w)$.

2. Per la seconda proprietà abbiamo:

$$Q(\alpha v) = \varphi(\alpha v, \alpha v) = \alpha^2 \varphi(v, v) = \alpha^2 Q(v).$$

□

Osserviamo che questo risultato mostra chiaramente che una forma quadratica NON è una forma lineare.

Proposizione 10.2.1. *Data una forma quadratica Q, la forma bilineare φ a cui essa è associata è ricostruibile nel modo seguente:*

$$\varphi(v,w) = \frac{1}{2}(Q(v+w) - Q(v) - Q(w)). \tag{10.3}$$

Dimostrazione. È una semplice conseguenza della proprietà 1 del Teorema 10.2.1. □

La (10.3) è detta *forma polare* della forma quadratica Q.

Un'altra ovvia conseguenza del Teorema 10.2.1 e di quest'ultima proposizione è il seguente.

Corollario 10.2.1. *La corrispondenza tra forme bilineari simmetriche e forme quadratiche associate è biunivoca.*

Definizione 10.2.2. La matrice associata a una forma quadratica Q rispetto a una base B è la matrice $M_B(\varphi) = (b_{i,j})_{i,j=1,\dots,n}$ associata alla forma bilineare simmetrica φ a cui Q è associata rispetto alla stessa base B:

$$Q(v) = \varphi(v,v) = \sum_{i,j=1}^{n} b_{i,j} v_i v_j = v^T M_B(\varphi) v.$$

Osservazione 10.2.1. Osserviamo che $Q(v)$ si esprime mediante un polinomio omogeneo di secondo grado nelle componenti del vettore v rispetto alla base B. Per questo motivo Q prende il nome di *forma quadratica*.

Definizione 10.2.3. Una forma quadratica si definisce *non degenere* se le matrici a essa associate sono di rango massimo. Si dice *rango di una forma quadratica* il rango di una delle sue matrici associate.

Definizione 10.2.4. Un vettore $v \in V$ si dice *isotropo* per una forma quadratica Q definita su V se $Q(v) = 0$.

Osserviamo che l'insieme $\mathcal{I}_Q$ dei vettori isotropi di una forma quadratica non è in generale un sottospazio vettoriale, infatti se $B = \{v_1, \dots, v_n\}$ è una base di V, esso può essere descritto come

$$\mathcal{I}_Q = \left\{ v = x_1 v_1 + \cdots + x_n v_n \mid \sum_{i=1}^{n} a_{i,j} x_i x_j = 0 \right\}$$

dove $(a_{i,j})$ è la matrice simmetrica associata a Q rispetto alla base B. L'insieme $\mathcal{I}_Q$ può quindi essere visto come il luogo degli zeri di un polinomio omogeneo di secondo grado. Se ricordiamo la definizione delle quadriche ci risulterà intuitivo perché tale insieme venga denominato *cono isotropo* di Q.

Definizione 10.2.5. Una forma quadratica Q su uno spazio vettoriale V si dice

- *definita positiva* (*negativa*) se

$$Q(v) \geq 0, (\leq 0),\ \forall v \in V \text{ e } Q(v) = 0 \iff v = 0,$$

- *semidefinita positiva* (*negativa*) se

$$Q(v) \geq 0, (\leq 0),\ \forall v \in V,$$

- *indefinita* altrimenti.

Definizione 10.2.6. Una forma quadratica Q su uno spazio vettoriale V si dice ridotta in *forma canonica* quando esiste una base B di V tale che per ogni $v = (x_1, \dots, x_n)_B \in V$ si abbia che

$$Q(v) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2.$$

Abbiamo visto nella prima parte del teorema di Sylvester della sezione precedente che per una forma bilineare simmetrica è sempre possibile trovare una base B tale che la sua matrice rappresentativa rispetto a tale base sia diagonale. Quindi, per definizione di forma quadratica, è sempre possibile trovare una base B di V tale che Q sia scrivibile in forma canonica. Questa cosa è spiegata più precisamente nel seguente teorema.

Teorema 10.2.2. *Sia V uno spazio vettoriale reale di dimensione n e B una sua base. Una forma quadratica Q su V ammette sempre una scrittura in forma canonica del tipo*

$$Q(v) = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

dove $\lambda_1, \ldots, \lambda_n$ sono gli autovalori di $M_B(Q)$ e $\{y_1, \ldots, y_n\}$ sono le coordinate di v rispetto a una base B' di autovettori di $M_B(Q)$.

Dimostrazione. La matrice $M_B(Q)$ è simmetrica, pertanto è diagonalizzabile a una matrice con gli autovalori sulla diagonale: esiste P matrice ortogonale tale che $P^{-1}M_B(Q)P = diag(\lambda_1, \ldots, \lambda_n) = D$. Ora facciamo il cambiamento di base indotto proprio da P:
$Q(v) = v'_B M_B(Q) v_B = (Pv_{B'})^T M_B(Q)(Pv_{B'}) =$
$= v_{B'}^T (P^T M_B(Q) P) v_{B'} = v_{B'}^T D v_{B'} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$ □

Nei seguenti esempi mostriamo due metodi alternativi per calcolare la forma canonica di una forma quadratica: il primo metodo consiste nell'applicare il teorema appena visto, il secondo metodo è il cosiddetto *metodo del completamento dei quadrati*.

Esempio 10.2.2. Consideriamo la seguente forma quadratica $Q : \mathbb{R}^2 \to \mathbb{R}$ così definita $Q((x, y)) = 3x^2 - 2xy + 3y^2$. La matrice associata a Q rispetto alla base canonica è

$$M_{\mathcal{E}}(Q) = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

i cui autovalori sono $\lambda_1 = 2, \lambda_2 = 4$ quindi la forma canonica associata a Q è

$$Q = 2y_1^2 + 4y_2^2$$

dove (y_1, y_2) sono le coordinate di un vettore rispetto alla base B che diagonalizza $M_{\mathcal{E}}(Q)$ ossia $B = \{(-1/\sqrt{2}, -1/\sqrt{2}), (-1/\sqrt{2}, 1/\sqrt{2})\}$.

La stessa forma canonica poteva essere trovata mediante il cosidetto *metodo del completamento dei quadrati*:
$Q = 3x^2 - 2xy + 3y^2 = (3x^2 - 2xy) + 3y^2 =$
$= (3x^2 - 2xy + 1/3y^2) + 3y^2 - 1/3y^2 =$
$= (\sqrt{3}x - 1/\sqrt{3}y)^2 + 8/3y^2$. Ora col cambiamento di base:

$$\begin{cases} y_1 = \sqrt{3/2}x - 1/\sqrt{6}y \\ y_2 = \sqrt{2/3}y \end{cases}$$

ritroviamo esattamente la forma canonica trovata col metodo degli autovalori (NB: la matrice associata, rispetto alla base canonica, al cambiamento di base appena effettuato ha determinante uguale a 1).

Esempio 10.2.3. Consideriamo la forma quadratica $Q : \mathbb{R}^3 \to \mathrm{R}$, con $Q((x, y, z)) = 2xy - 2yz$. La matrice associata a Q rispetto alla base canonica è

$$M_{\mathcal{E}}(Q) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

i cui autovalori sono $\lambda_1 = \sqrt{2}, \lambda_2 = -\sqrt{2}, \lambda_3 = 0$. Quindi la forma canonica di Q è

$$Q = \sqrt{2}y_1^2 - \sqrt{2}y_2^2$$

dove (y_1, y_2, y_3) sono lo coordinate di un vettore di $\mathbb{R}^3$ scritto rispetto alla base B che diagonalizza $M_{\mathcal{E}}(Q)$, ossia una base di autovettori. Calcolando gli autovettori di $M_{\mathcal{E}}(Q)$ si trova che la base B cercata è $\{(-1/2, -1/\sqrt{2}, 1/2), (-1/2, 1/\sqrt{2}, 1/2), (1/\sqrt{2}, 0, 1/\sqrt{2})\}$.

La stessa forma canonica poteva essere trovata nuovamente con il metodo del completamento dei quadrati ma questa volta applicato in modo meno ovvio in quanto nell'espressione che abbiamo di Q mancano i termini al quadrato: $Q = 2xy - 2yz = 2y(x - z)$. Facciamo dapprima un cambio di variabili $(x - z) \to \sqrt{2}X$, $y \to Y$ e $z \to Z$ da cui otteniamo che $Q = 2\sqrt{2}XY$. Ora per ricavare i termini al quadrato occorre fare un altro cambio di variabili: $X \to (1/\sqrt{2})(X' + Y')$ e $Y \to (1/\sqrt{2})(X' - Y')$ da cui $Q = \sqrt{2}(X'^2 - Y'^2)$ che è proprio la forma canonica di cui sopra. (NB: le matrici associate, rispetto alla base canonica, a tutti i cambiamenti di base che abbiamo effettuato hanno determinante uguale a 1).

In questo esempio abbiamo visto un caso particolare del seguente teorema.

Teorema 10.2.3 (Teorema di Gauss, Lagrange). *Sia Q una forma quadratica su uno spazio vettoriale V, mediante il metodo del completamento dei quadrati è possibile determinare una base di V che permette di scrivere la forma quadratica Q in forma canonica.*

Sempre il teorema di Sylvester suggerisce che una forma quadratica possa essere scritta in forma ancora più essenziale.

Definizione 10.2.7. Una forma quadratica Q su uno spazio vettoriale V si dice ridotta in *forma normale* se è ridotta a forma canonica e i coefficienti del polinomio omogeneo che ne deriva sono $1, -1, 0$.

Teorema 10.2.4. *Una forma quadratica Q definita su uno spazio vettoriale V di dimensione n, ammette sempre una forma normale. Sia B una base di V, e sia $Q(v) = v^T M_B(Q) v$. Il numero dei coefficienti uguali a 1 (-1 o 0 rispettivamente), che compare nella forma normale, è uguale al numero di autovalori positivi (negativi o nulli rispettivamente) della matrice $M_B(Q)$, contati con la relativa molteplicità.*

Dimostrazione. Abbiamo già visto che una forma quadratica ammette sempre una scrittura in forma canonica. Sia $B' = \{v_1, \dots, v_n\}$ la base che permette di scrivere Q in forma canonica. Ora, è ovvio che, eventualmente riordinando i vettori della base B', esistono due interi $p, q \geq 0$ tali che $Q(v_i) > 0$ per $i = 1, \dots, p$, $Q(v_j) < 0$ per $j = p+1, \dots, p+q$ e $Q(v_k) = 0$ per i rimanenti $k = p+q+1, \dots, n$. Come nel caso della dimostrazione del teorema di Sylvester, è sufficiente considerare la base $\tilde{B} = \left\{ \frac{v_1}{\sqrt{Q(v_1)}}, \dots, \frac{v_p}{\sqrt{Q(v_p)}}, \frac{v_{p+1}}{\sqrt{-Q(v_{p+1})}}, \dots, \frac{v_{p+q}}{\sqrt{-Q(v_{p+q})}}, v_{p+q+1}, \dots, v_n \right\}$ per ottenere che $M_{\tilde{B}}(Q) = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$. □

Osservazione 10.2.2. Se applichiamo il precedente teorema al caso in cui $V = \mathbb{R}^3$, troveremo la classificazione delle superfici quadriche, mentre se invece $V = \mathbb{R}^2$ avremo la classificazione delle coniche (vedi la prossima sezione).

Corollario 10.2.2. *Una forma quadratica ammette una sola forma normale.*

Per calcolare la forma normale di una forma quadratica si può quindi o calcolare dapprima la sua forma canonica e poi ridurre a ± 1 tutti i suoi coefficienti mantenendo il segno che avevano nella forma canonica, oppure si può applicare la cosidetta regola dei segni di Cartesio.

Teorema 10.2.5 (Regola dei segni di Cartesio)**.** *Dato un polinomio a coefficienti reali $f = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$, avente tutte le radici reali, si ha che le radici positive di f sono tante quante sono le variazioni di segno della successione $a_0, \dots, a_n$.*

Questa regola si applica al polinomio caratteristico di una qualsiasi matrice associata a una forma quadratica. Essendo tale matrice simmetrica, esso avrà tutte le sue radici reali. Quindi per classificare una forma quadratica Q si può procedere anche nel modo seguente:

1. Si trova una matrice A, $n \times n$, associata all forma quadratica Q (non importa rispetto a quale base);
2. Se ne calcola il polinomio caratteristico $P_A(\lambda) = \lambda^s R(\lambda)$;
3. La molteplicità s dell'autovalore nullo coincide con la dimensione del nucleo di A, ossia $s = n - rk(A)$;
4. Si calcola il numero p delle variazioni di segno del polinomio $R(\lambda)$ ed esso sarà il numero delle radici positive;
5. Il numero delle radici negarive si calcola ora semplciemente facendo $n - s - p$ in quanto una matrice simmetrica ha tutti gli autovalori reali.
6. La forma quadratica avrà quindi forma normale

$$x_1^2 + \cdots x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2.$$

Un'altro aspetto delle forme quadratiche è che possono anche essere viste come una generalizzazione del concetto di "norma". Abbiamo visto come la definizione di prodotto scalare coincida con quella di forma bilineare simmetrica definita positiva. Ora sia $\varphi : V \times V \to \mathbb{R}$ una forma bilineare definita positiva (e quindi un prodotto scalare), se rileggiamo in questo nuovo linguaggio la definizione di norma possiamo dire che la norma di un vettore $v \in V$ rispetto a φ si definisce come $||v||_\varphi := \sqrt{\varphi(v,v)}$, ossia

$$||v||_\varphi = \sqrt{\varphi(v,v)} = \sqrt{Q(v)}.$$

In questo senso le forme quadratiche possono essere viste come una definizione alternativa della norma rispetto a un prodotto scalare associato a una forma bilineare simmetrica definita positiva. In quest'ottica possiamo rileggere anche i concetti di ortogonalità e ortonormalità.

Definizione 10.2.8. Sia V uno spazio vettoriale di dimensione n su cui è definita una forma bilineare simmetrica φ associata a una forma quadratica Q. Una base $B = (v_1, \ldots, v_n)$ si dice *ortogonale* rispetto a φ se:

$$\varphi(v_i, v_j) = 0,\ i, j = 1, \ldots, n,\ i \neq j.$$

Una base $B = (e_1, \ldots, e_n)$ si dice *ortonormale* rispetto a φ se è ortogonale rispetto a φ e se

$$Q(v_i) = 1,\ i = 1, \ldots, n.$$

Chiaramente ogni base di V che permette di scrivere una forma quadratica in forma canonica è una base ortogonale, quindi almeno una base ortogonale esiste sempre.

Inoltre dal Teorema 10.2.4 segue anche che solo le forme quadratiche la cui forma normale è quella avente tutti i coefficienti uguali a 1 ammettono basi ortonormali. È possibile mostrare che esse sono necessariamente dei prodotti scalari.

10.3 Coniche

Questa sezione è indipendente dalle precedenti del presente capitolo: la trattazione che diamo delle coniche prescinde dalla conoscenza delle forme bilineari e delle forme quadratiche. Ciononostante il lettore che ha già studiato le due sezioni precedenti ritroverà qui i richiami necessari per comprendere le coniche come caso particolare di forme quadratiche.

10.3.1 Introduzione

Le curve piane dette "sezioni coniche" sono state studiate dai matematici fin dall'antichità. Le prime tracce del loro studio risalgono al matematico greco Menecmo (375-325 a.C. circa).

Le coniche furono dapprima definite come l'intersezione fra un cono circolare retto (a due falde) con angolo al vertice variabile e un piano perpendicolare alla generatrice del cono. In dipendenza del fatto che l'angolo fosse minore, uguale o maggiore dell'angolo retto si ottenevano figure diverse. Si è poi affermato il punto di vista di considerare il cono con angolo al vertice fisso (retto), e variare la posizione del piano in modo da ottenere rispettivamente un'ellisse (vedi Figure 10.1 e 10.2), una parabola (Figura 10.3) o un'iperbole (Figura 10.4).

Nelle seguenti figure si interseca un cono circolare a due falde con piani aventi varia inclinazione. Ricordiamo che si dice *asse* del cono la retta (per il vertice) che è il suo asse di simmetria e si dice *direttrice* del cono una qualsiasi retta in esso contenuta e passante per il suo vertice (il cono può essere visto come la superficie di rotazione ottenuta facendo ruotare una direttrice di 360° attorno all'asse).

Figura 10.1: Circonferenza: Sezione conica ottenuta intersecando un cono circolare con un piano ortogonale all'asse del cono (una circonferenza è una speciale ellisse).

Figura 10.2: Ellisse: Sezione conica ottenuta intersecando un cono circolare con un piano che forma con l'asse un angolo maggiore di quello della direttrice.

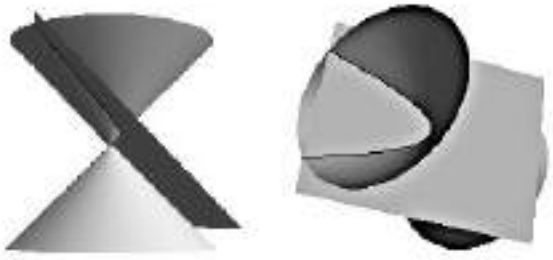


Figura 10.3: Parabola: Sezione conica ottenuta intersecando un cono circolare con un piano parallelo a una direttrice.

Figura 10.4: Iperbole: Sezione conica ottenuta intersecando il cono quadrico con un piano che forma con l'asse un angolo minore di quello della direttrice. Si nota che l'iperbole è formata da due componenti dette *falde*

Apollonio di Perga (262-190 a.c. circa) consolidò i risultati fino a quel momento trovati nella sua opera *Le Coniche*, costituita da otto libri (di alcuni è arrivato sino a noi l'originale greco, di altri la traduzione araba, e l'ottavo è andato perduto ma se ne conosce il contenuto attraverso citazioni e ricostruzioni). Apollonio usò sezioni piane di un unico cono circolare con piani di diversa inclinazione. Molti altri matematici nella storia si sono dedicati a questo studio e ne hanno utilizzato e perfezionato i risultati; tra questi ricordiamo Keplero, Cartesio, Fermat, Newton.

Le coniche possono essere definite in due modi del tutto equivalenti. Il primo, con cui iniziamo, dà la loro descrizione come luogo geometrico di punti del piano (Sezione 10.3.2). Nella Sezione 10.3.3 ne generalizzeremo la definizione algebrica tramite un'equazione di secondo grado in due variabli.

10.3.2 Le coniche come luoghi geometrici

Definizione 10.3.1. Fissati una retta $\mathcal{D}$ del piano $\mathbb{R}^2$ (*Direttrice*) e un punto $F \in \mathbb{R}^2$ (*Fuoco*) tale che F non appartenga alla retta $\mathcal{D}$, si definisce *conica* $\mathcal{C}$ il luogo geometrico dei punti $P \in \mathbb{R}^2$ tali che il rapporto tra la distanza di P da F e di P da $\mathcal{D}$ sia costante. Tale costante è chiamata *eccentricità* e viene usualmente indicata con la lettera e.

Vediamo concretamente questa costruzione. Sia H il punto della retta $\mathcal{D}$ ottenuto come intersezione tra $\mathcal{D}$ e la retta passante per un punto P della conica e perpendicolare a $\mathcal{D}$ (vedi Figura 10.5) allora l'eccentricità e è così ottenuta:

$$\frac{\overline{PF}}{\overline{PH}} = e. \tag{10.4}$$

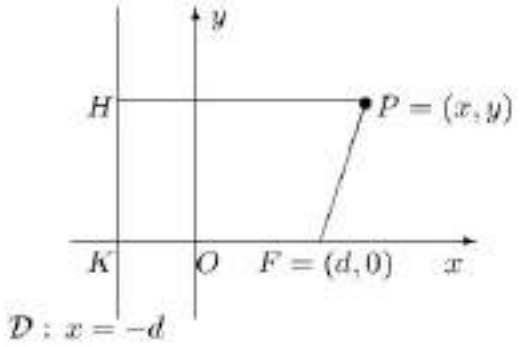


Figura 10.5: Costruzione di una conica come luogo geometrico di punti.

Scegliamo ora un sistema di riferimento cartesiano in modo tale che l'asse x sia ortogonale a $\mathcal{D}$ e passi per F. Sia K il punto di intersezione tra l'asse x e la retta $\mathcal{D}$, e l'asse y intersechi l'asse x nel punto medio di $\overline{KF}$ (Figura 10.5). In questo sistema di riferimento supponiamo che il punto F abbia coordinate $(d, 0)$, cosicché la retta $\mathcal{D}$ avrà equazione $x = -d$. Se assumiamo che il punto P abbia le generiche coordiante (x, y), allora la (10.4) diventa:

$$\frac{\sqrt{y^2 + (x-d)^2}}{x+d} = e \tag{10.5}$$

Osservazione 10.3.1. La (10.5) è ben definita se il denominatore è non nullo; ciò equivale a richiedere $x \neq -d$ ossia che il punto P non appartenga mai alla retta $\mathcal{D}$. Ciò significa che l'intersezione delle nostre coniche con la direttrice sarà sempre vuota.

Osservazione 10.3.2. L'eccentricità, trattandosi di rapporto di distanze tra punti, non potrà mai assumere valori negativi.

Dalla (10.5) si ottiene facilmente l'equazione generale delle coniche:

$$(1-e^2)x^2 + y^2 - 2d(e^2+1)x + d^2(1-e^2) = 0. \tag{10.6}$$

Avremo i vari tipi di conica, a seconda del valore assegnato all'eccentricità e:

Definizione 10.3.2. Una conica definita come in 10.3.1 si dice un'*ellisse* se $e < 1$; una *parabola* se $e = 1$ e un' *iperbole* se $e > 1$.

Distinguiamo i vari casi e studiamoli separatamente: $e = 1$, $e > 1$ ed $e < 1$.

1. Se $e = 1$ allora $(1 - e^2) = 0$ perciò la (10.6) diventa $y^2 - 4dx = 0$, da cui la seguente *equazione canonica della Parabola*:

$$x = \frac{1}{4d}y^2. \tag{10.7}$$

Richiedere che e sia uguale a 1 equivale a richiedere che $\overline{PF} = \overline{PH}$ (vedi (10.4)). Se rileggiamo la definizione ci accorgiamo che i punti di una conica $\mathcal{C}$ con eccentricità uguale a 1 sono equidistanti da un punto e una retta fissati. Questa è precisamente la definizione della parabola come luogo geometrico, dunque la (10.7) è esattamente l'equazione canonica della *Parabola*. Il punto fisso F è detto *fuoco* e la retta fissa $\mathcal{D}$ è detta *direttrice*.

2. Supponiamo ora che sia $e \neq 1$. Ciò significa che $(1-e^2) \neq 0$. Quindi il termine $\frac{d^2(1+e^2)^2}{1-e^2}$ è ben definito. Aggiungiamolo e togliamolo alla (10.6); dopo un semplice raccoglimento, avremo:

$$(1-e^2)\left[x^2 - \frac{2d(e^2+1)}{1-e^2}x + \frac{d^2(1+e^2)^2}{1-e^2}\right] +$$

$$+y^2 + d^2(1-e^2) - \frac{d^2(1+e^2)^2}{1-e^2} = 0$$

da cui:

$$(1-e^2)\left(x - \frac{d(1+e^2)}{1-e^2}\right)^2 + y^2 - \frac{4d^2e^2}{1-e^2} = 0.$$

Operiamo ora il seguente cambiamento di coordinate (vedi Figura 10.6):

$$\begin{cases} X = x - \frac{d(1+e^2)}{1-e^2} \\ Y = y \end{cases} \tag{10.8}$$

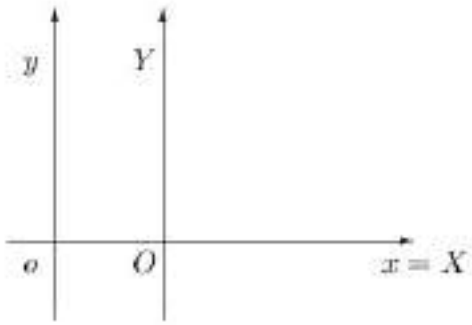


Figura 10.6: Traslazione definita in 10.8: si passa dal sistema di riferimento $\{oxy\}$ a $\{OXY\}$.

Osserviamo che ciò che abbiamo fatto, geometricamente, è di tenere fisso l'asse delle x e traslare l'asse delle y esattamente di $\frac{d(1+e^2)}{1-e^2}$. Con questo cambiamento di coordinate otteniamo:

$$(1-e^2)X^2 + Y^2 - \frac{4d^2e^2}{1-e^2} = 0$$

da cui

$$\frac{(1-e^2)^2}{4d^2e^2}X^2 + \frac{1-e^2}{4d^2e^2}Y^2 = 1.$$

Poniamo $\frac{1}{a^2} = \frac{(1-e^2)^2}{4d^2e^2}$ e otteniamo:

$$\frac{X^2}{a^2} + \left(\frac{1-e^2}{4d^2e^2}\right)Y^2 = 1. \tag{10.9}$$

Osserviamo ora il coefficiente di Y^2: il denominatore è sempre positivo, mentre il numeratore è positivo o negativo quando $e < 1$ oppure, rispettivamente, $e > 1$.

Si distinguono dunque i due casi.

3. Sia $e < 1$. Abbiamo appena osservato che da ciò si deduce che $\frac{1-e^2}{4d^2e^2} > 0$, è quindi lecito porre $\frac{1}{b^2} = \frac{1-e^2}{4d^2e^2}$ e ottenere così dalla (10.9) l'*equazione canonica dell'Ellisse*:

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1. \tag{10.10}$$

4. Sia ora $e > 1$ si ha che $\frac{1-e^2}{4d^2e^2} < 0$. Poniamo dunque $\frac{1}{b^2} = -\frac{1-e^2}{4d^2e^2}$ da cui la (10.9) diventa:

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1 \tag{10.11}$$

che è *l'equazione canonica dell'Iperbole*.

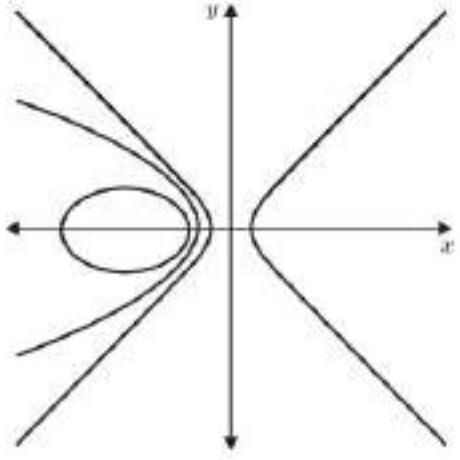


Figura 10.7: La figura mostra tre coniche di eccentricità rispettivamente $\sqrt{5}/3$ (ellisse), 1 (parabola) e $\sqrt{2}$ (iperbole).

10.3.3 Classificazione

Lo scopo di questa sezione è di classificare le coniche a partire dalla loro equazione rappresentativa (non ci inoltreremo nei dettagli delle dimostrazioni).

Osserviamo che l'equazione (10.6) che rappresenta tutti i tipi di coniche è un'equazione di secondo grado nelle variabili x e y. Ciò ci porta a generalizzare la definizione di che cosa sia una conica nel seguente modo:

Definizione 10.3.3. Dato un polinomio di secondo grado $f \in \mathbb{R}[x, y]$, il sottoinsieme $\mathcal{C} \in \mathbb{R}^2$ costituito dai punti le cui coordinate soddisfano l'equazione

$$f(x, y) = 0$$

è detto *una conica*.

Osservazione 10.3.3. Nella Definizione 10.3.3 si avranno, oltre alle coniche già viste nella sezione precedente (definite come in 10.6) anche altre "coniche" che costituiranno dei casi "degeneri" di quelle già viste. Si può ad esempio avere il caso in cui $\mathcal{C} = \emptyset$, se $f(x, y) = x^2 + y^2 + 1$; se invece $f(x, y) = x^2 + y^2$ avremo che la conica corrispondente è costituita dal solo punto $(0, 0)$, mentre se $f(x, y) = x^2 - y^2 = (x + y)(x - y)$ la conica è data da due rette.

Osservazione 10.3.4. Se $f(x, y) = 0$ definisce la conica $\mathcal{C}$ e se g è un altro polinomio di $\mathbb{R}[x, y]$ tale che $g = \lambda f$, con $\lambda \in \mathbb{R}\backslash\{0\}$, allora g e f definiscono la medesima conica $\mathcal{C}$.

Osservazione 10.3.5. Nel caso in cui il polinomio f della Definizione 10.3.3 sia omogeneo, la definizione di conica è equivalente a quella di forma quadratica (Definizione 10.2.1) sullo spazio vettoriale $\mathbb{R}^2$.

Supponiamo ora di trovarci davanti a un polinomio di secondo grado di $\mathbb{R}[x, y]$:

$$\alpha x^2 + \beta xy + \gamma y^2 + \delta x + \epsilon y + \phi = 0 \tag{10.12}$$

dove $\alpha, \beta, \gamma, \delta, \epsilon, \phi \in \mathbb{R}$ e α, β, γ non sono tutti nulli.

Per la definizione 10.3.3, abbiamo a che fare con una conica; per stabilire di quale conica si tratti le operazioni geometriche che si fanno sono una traslazione e una rotazione del sistema di riferimento cartesiano, in modo che esso venga a coincidere con quello indicato nella sezione precedente con le lettere X, Y[1]; quest'operazione equivale a trasformare l'equazione (10.12) nella sua cosiddetta *forma canonica*, come abbiamo già visto per parabola, ellisse e iperbole.

[1] L'operazione di rotazione del sistema di riferimento è proprio quella che permette di far scomparire nell'equazione in forma canonica il termine $x \cdot y$.

Proposizione 10.3.1. *Ogni conica può essere portata, tramite un opportuno cambiamento di coordinate, a essere rappresentata da un'equazione di quelle nell'elenco seguente, dette forme canoniche. Nel seguito si ha* $a, b \in \mathbb{R}$, $a > 0$, $b > 0$.

- **Ellisse reale:** *(Fig. 10.8).*
$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1.$$
- **Ellisse a punti non reali:**
$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = -1.$$
Le soluzioni dell'equazione sono soltanto in $\mathbb{C}^2$.
- **Ellisse degenere:**
$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 0.$$
Possiede solo un punto, l'origine; dal punto di vista complesso si tratta di due rette: $Y = \pm\frac{biX}{a}$.
- **Iperbole reale:** *(Fig. 10.9)*
$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1.$$
- **Iperbole degenere:** *(Fig. 10.9).*
$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 0.$$
Sono due rette: $Y = \pm\frac{bX}{a}$.
- **Parabola:** *(Fig. 10.10)*
$$Y^2 - 2pX = 0, \quad p > 0.$$
- **Parabola degenere:** *(Fig. 10.10).*
$$Y^2 - a^2 = 0.$$
Sono due rette reali parallele $Y = \pm a$.
- **Parabola degenere** $= \emptyset$*:*
$$Y^2 + a^2 = 0.$$
Sarebbero due rette immaginarie $Y = \pm ai$.
- **Conica doppiamente degenere:**
$$Y^2 = 0.$$
È la retta $Y = 0$*, doppia.*

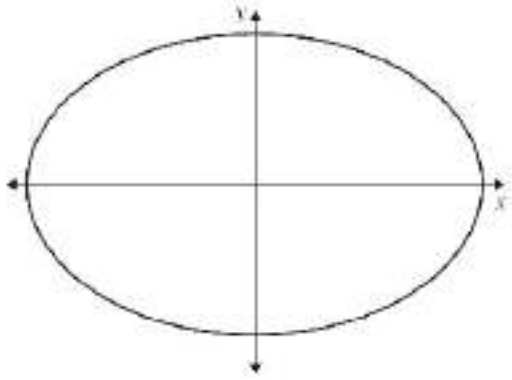


Figura 10.8: Ellisse reale.

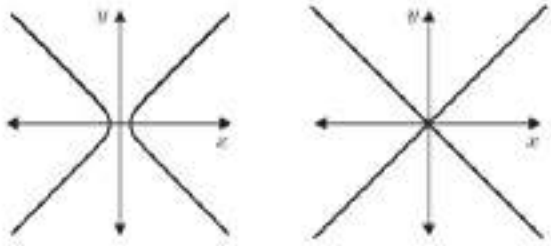


Figura 10.9: Iperbole reale a sinistra, iperbole degenere in due rette a destra.

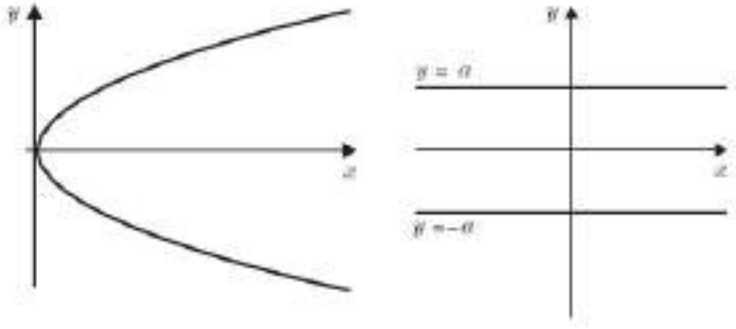


Figura 10.10: Parabola a sinistra, parabola degenere in due rette a destra.

Un metodo immediato per riconoscere se la (10.12) rappresenta un'ellisse, una parabola o un'iperbole è il seguente. Indichiamo con

$$\Delta := \beta^2 - 4\alpha\gamma$$

allora

$$\begin{cases} \text{se } \Delta < 0 & \Rightarrow \mathcal{C} \text{ è un'ellisse;} \\ \text{se } \Delta = 0 & \Rightarrow \mathcal{C} \text{ è una parabola;} \\ \text{se } \Delta > 0 & \Rightarrow \mathcal{C} \text{ è un'iperbole.} \end{cases}$$

Il termine $\beta^2 - 4\alpha\gamma = \Delta$ porta il nome di *discriminante* della curva, e vedremo in che modo è collegato alla forma della conica.

Osservazione 10.3.6. Abbiamo già notato nell'Osservazione 10.3.5 un legame tra coniche e forme quadratiche nel caso in cui il polinomio che definisce la conica sia omogeneo. Se vogliamo rileggere la classificazione delle coniche in termini di forme quadratiche quello che dobbiamo fare è la cosa seguente: omogeneizziamo il polinomio di definizione della conica con una terza variabile z in modo che tutti i monomi siano di grado 2 (ad esempio se $f = x^2 + xy + x - 2$ allora consideriamo il polinomio $\tilde{f} = x^2 + xy + xz - 2z^2$); in questo modo abbiamo un polinomio di grado 2 in tre variabili e può quindi essere considerato una forma quadratica su $\mathbb{R}^3$. Se ora scriviamo la forma canonica di $\tilde{f}$ come si è fatto nella Sezione 10.2 ritroviamo che essa coinciderà con una delle forme canoniche della Proposizione 10.3.1 modulo porre $z = 1$.

10.3.4 Coniche e isometrie

Come si è già osservato una conica $\mathcal{C}$ può essere scritta nella forma (10.12) che possiamo anche riscrivere come segue:

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0 \tag{10.13}$$

con $a, b, c, d, e, f \in \mathbb{R}$ e a, b, c non simultaneamente nulli.

Consideriamo la matrice simmetrica

$$A = \begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix} \tag{10.14}$$

L'equazione (10.13) può anche essere rappresentata nella seguente forma matriciale:

$$\begin{pmatrix} x & y & 1 \end{pmatrix} \cdot A \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0. \tag{10.15}$$

Consideriamo ora un cambiamento di coordinate del tipo:

$$\begin{pmatrix} x \\ y \end{pmatrix} = M_0 \cdot \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{10.16}$$

dove $M_0 = (m_{ij}) \in \mathbb{R}^{2,2}$ è una matrice ortogonale. In dettaglio si ha:

$$\begin{cases} x = m_{11}X + m_{12}Y + c_1 \\ y = m_{21}X + m_{22}Y + c_2 \end{cases}$$

Un cambiamento di coordinate di questo tipo NON è un'applicazione lineare (ad esempio l'origine viene portata nel punto (c_1, c_2)) ed è composto da un'isometria lineare $\begin{pmatrix} x \\ y \end{pmatrix} = M_0 \cdot \begin{pmatrix} X \\ Y \end{pmatrix}$ (che è un'applicazione lineare) e da una traslazione (non lineare) $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$. È una trasformazione che conserva distanze e angoli in quanto la matrice M è ortogonale, quindi conserva il prodotto scalare, e una traslazione è un movimento rigido del piano che non altera angoli e distanze. Tale cambiamento di coordinate viene anche chiamato *isometria*, generalizzando quanto visto nella Sezione 8.4 per le applicazioni lineari.

Non è difficile persuaderci del fatto che se operiamo sulla conica $\mathcal{C}$ di equazione (10.13) con una trasformazione come l'isometria appena vista, quello che otterremo è una conica $\mathcal{C}'$ che ha la stessa forma di $\mathcal{C}$, praticamente $\mathcal{C}'$ è la stessa conica, solo "spostata" (traslata) e/o ruotata o "ribaltata" (vedi ancora Sezione 8.4).

Per meglio rappresentare l'equazione di $\mathcal{C}'$ esprimiamo l'isometria in modo matriciale:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & c_1 \\ m_{21} & m_{22} & c_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}. \tag{10.17}$$

Che ci dà anche (trasponendo e ricordando che ${}^tAB = {}^tB{}^tA$):

$$(x\ y\ 1) = (X\ Y\ 1) \cdot {}^tM.$$

Eseguendo queste sostituzioni nella (10.15) otteniamo l'equazione di $\mathcal{C}'$ nelle nuove variabili X, Y:

$$\begin{pmatrix} X & Y & 1 \end{pmatrix} B \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

dove $B = {}^tMAM$ e M è proprio la matrice che incontriamo in (10.17). Dalla relazione $B = {}^tMAM$ si ha che $rg(A) = rg(B)$, da cui si deduce la seguente Proposizione.

Proposizione 10.3.2. *Il rango della matrice A associata a una conica $\mathcal{C}$ è tale che tutte le coniche ottenute da essa tramite un cambiamento di variabili che sia un'isometria hanno il medesimo rango della matrice associata. Esso è chiamato* rango di $\mathcal{C}$ *e denotato con $r(\mathcal{C})$.*

Osservazione 10.3.7. Il lettore che abbia studiato anche la Sezione 10.2 avrà riconosciuto in questo procedimento il metodo del completamento dei quadrati descritto nel Teorema 10.2.3 di Gauss e Lagrange.

Possiamo quindi dare la seguente Definizione:

Definizione 10.3.4. Una conica si dice *non degenere, degenere, semplicemente degenere, doppiamente degenere* se si ha, rispettivamente, che $r(\mathcal{C}) = 3$, $r(\mathcal{C}) < 3$, $r(\mathcal{C}) = 2$, $r(\mathcal{C}) = 1$.

Siano ora A_0, B_0 e M_0 le sottomatrici ottenute rispettivamente da A, da B e da M eliminando la terza riga e la terza colonna. Si verifica semplicemente che vale anche:

$$B_0 = {}^tM_0A_0M_0.$$

Ciò implica che A_0 e B_0 hanno lo stesso rango, e che il segno di $\det(A_0)$ è lo stesso di $\det(B_0)$ (usando il teorema di Binet), da cui si ha la seguente Proposizione:

Proposizione 10.3.3. *Il rango di A_0 e il segno del suo determinante sono proprietà comuni a ogni conica ottenuta da $\mathcal{C}$ tramite una isometria.*

Si ha poi che

Definizione 10.3.5. Se $\det(A_0) \neq 0$, $\mathcal{C}$ si dice una *conica a centro*; altrimenti se $\det(A_0) = 0$, $\mathcal{C}$ è una parabola.

E che

Proposizione 10.3.4. *Una conica a centro con $\det(A_0) > 0$ è un'ellisse; una conica a centro con $\det(A_0) < 0$ è un'iperbole.*

Quest'ultima proposizione giustifica la classificazione tramite il discriminante (infatti $-\Delta = 4\det(A_0)$.), in effetti, i principali strumenti di cui ci avverremo per classificare una conica saranno il determinante di $A_{\mathcal{C}}$ e quello di A_0:

- Se il determinante di $A_{\mathcal{C}} \neq 0$ la conica si dice "non degenere",

- Se il determinante di $A_{\mathcal{C}} = 0$ la conica si dice "degenere". In particolare:
  - Se $rg(A_{\mathcal{C}}) = 2$ la conica è "semplicemente degenere".
  - Se $rg(A_{\mathcal{C}}) = 1$ la conica è "doppiamente degenere".

Inoltre, notiamo che scegliendo opportunamente la matrice M_0, avremo che $B_0 = {}^tMAM$ sarà una matrice diagonale, infatti A_0, essendo simmetrica, è ortogonalmente diagonalizzabile (dal teorema spettrale) e quindi nell'equazione di $\mathcal{C}'$ non compare il termine in XY e, scegliendo opportunamente anche la traslazione (c_1, c_2), l'equazione della conica può esser riportata in una delle *forme canoniche* viste della Proposizione 10.3.1.

- Se $(\det(A_0)) \neq 0$ allora $\mathcal{C}$ è una conica a centro. Infatti si ha $\det A_0 = \det B_0 \neq 0$, che implica che nella forma canonica i coefficienti di X^2 e Y^2 non sono nulli. In particolare
  - Se $\det(A_0) > 0$ allora $\mathcal{C}$ è un'ellisse.
  - Se $\det(A_0) < 0$ allora $\mathcal{C}$ è un'iperbole.
- Se $\det(A_0) = 0$ allora $\mathcal{C}$ è una conica non a centro ossia una parabola. Infatti in questo caso un autovalore di A_0 è $\lambda = 0$, quindi B_0 ha uno zero sulla diagonale e nell'equazione di $\mathcal{C}'$ appare solo X^2 (o Y^2).

10.4 Quadriche

Un argomento completamente analogo a quello appena trattato in $\mathbb{R}^2$ per classificare le coniche può essere ripetuto in $\mathbb{R}^3$, ovviamente con le opportune modifiche, per classificare le superficie quadriche.

Definizione 10.4.1. Il luogo geometrico $\mathcal{Q}$ dei punti $(x, y, z) \in \mathbb{R}^3$ che soddisfano una equazione di secondo grado:

$$ax^2 + 2bxy + cxz + dy^2 + 2eyz + fz^2 + 2gx + 2hy + 2iz + l = 0$$

è chiamato *superficie quadrica*.

L'equazione descrivente $\mathcal{Q}$ può essere scritta in forma matriciale:

$$\begin{pmatrix} x & y & z & 1 \end{pmatrix} A_{\mathcal{Q}} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

dove

$$A_{\mathcal{Q}} = \begin{pmatrix} a & b & c & g \\ b & d & e & h \\ c & e & f & i \\ g & h & i & l \end{pmatrix}.$$

La sottomatrice $A_0 \in \mathbb{R}^{3,3}$ che si ottiene eliminando da $A_{\mathcal{Q}}$ la quarta riga e la quarta colonna avrà una funzione analoga per la classificazione di $\mathcal{Q}$ di quella avuta da $A_0 \in \mathbb{R}^{2,2}$ per le coniche.

10.4.1 Classificazione delle quadriche

Per classificare le quadriche terremo conto del determinante di $A_{\mathcal{Q}}$ e del determinante di A_0, in analogia con quanto visto per le coniche. Anche in questo caso avremo che con un cambiamento di coordinate che sia un'isometria, composto da un'isometria lineare (definita da una matrice ortogonale) e da una traslazione, che potremo rappresentare di nuovo con un'unica matrice M, e potremo portare $\mathcal{Q}$ in una forma $\mathcal{Q}'$ data da una matrice $B = {}^tMA_{\mathcal{Q}}M$ tale che $B_0 \in \mathbb{R}^{3,3}$ sia diagonale, e quindi ricondurci a una *forma canonica* della quadrica, e classificare in questo modo le possibili quadriche. Avremo allora:

- Se $\det(A_{\mathcal{Q}}) \neq 0$ la quadrica sarà non degenere.
  - Se $\det(A_0) \neq 0$ la quadrica sarà "a centro";
  - Se $\det(A_0) = 0$ la quadrica sarà "non a centro";
- Se $\det(A_{\mathcal{Q}}) = 0$ la quadrica sarà degenere.

Più in dettaglio avremo la seguente classificazione:

Proposizione 10.4.1. *Le possibili forme canoniche delle quadriche sono (qui $a, b, c > 0$):*

- *Quadriche non a centro non degeneri:*
  - *paraboloide ellittico:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} - z = 0$,
  - *paraboloide iperbolico (a sella):* $\frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0$.
- *Quadriche a centro non degeneri:*
  - *ellissoide:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0$,
  - *iperboloide a una falda (iperbolico):* $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0$,

  - *iperboloide a due falde (ellittico):* $\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0$,
  - *ellissoide non reale:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} + 1 = 0$ *(*$\mathcal{Q} = \emptyset$*).*

- *Quadriche degeneri con* $r(A_\mathcal{Q}) = 3$*:*
  - *Cono ellittico:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} - z^2 = 0$,
  - *Cono ellittico immaginario:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} + z^2 = 0$ *(*$\mathcal{Q} = \{(0,0,0)\}$*)*,
  - *Cilindro ellittico:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$,
  - *Cilindro ellittico immaginario:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$ *(*$\mathcal{Q} = \emptyset$*)*,
  - *Cilindro parabolico:* $x^2 - 2py = 0$, $p > 0$,
  - *Cilindro iperbolico:* $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$.

- *Quadriche degeneri con* $r(A_\mathcal{Q}) = 2$*:*
  - *Due piani incidenti:* $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$,
  - *Due piani complessi incidenti:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$ *(*$\mathcal{Q}$ = *asse* z*)*,
  - *Due piani paralleli:* $x^2 - p = 0$, $p > 0$,
  - *Due piani complessi paralleli:* $x^2 + p = 0$, $p > 0$.

- *Quadriche degeneri con* $r(A_\mathcal{Q}) = 1$*:*
  - *Un piano doppio:* $x^2 = 0$.

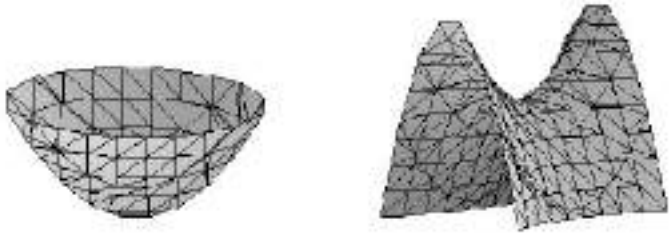


Figura 10.11: Quadriche non a centro non degeneri. Da sinistra: Paraboloide ellittico, paraboloide iperbolico (a sella).

Osservazione 10.4.1. Così come nell'Osservazione 10.3.6 si è visto come recuperare la classificazione delle coniche tramite la classificazione delle forme quadratiche di $\mathbb{R}^3$, una costruzione del tutto analoga può essere fatta per classificare le quadriche usando la forma canonica di una forma quadratica di $\mathbb{R}^4$. Omogeneizziamo il polinomio f di definizio-

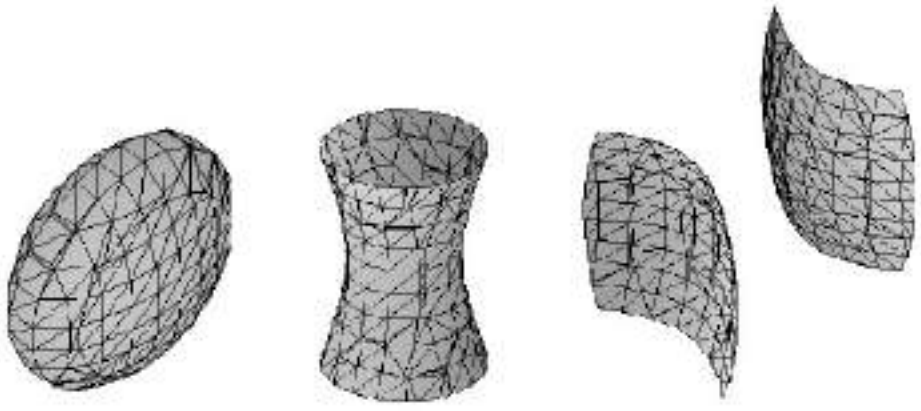


Figura 10.12: Quadriche a centro non degeneri reali. Da sinistra: Ellissoide, iperboloide a una falda (iperbolico), iperboloide a due falde (ellittico).

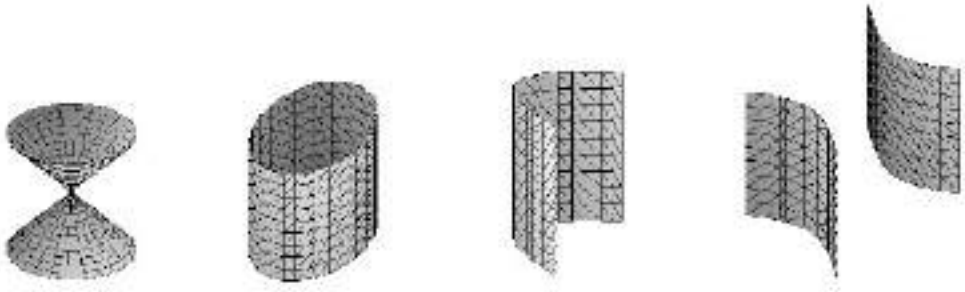


Figura 10.13: Quadriche degeneri con $r(A_Q) = 3$. Da sinistra: Cono ellittico, cilindro ellittico, cilindro parabolico, cilindro iperbolico.

ne della quadrica con una quarta variabile t e otteniamo un polinomio omogeneo di grado 2 in quattro variabili, ossia una forma quadratica di $\mathbb{R}^4$; scriviamo la sua forma canonica; deomgeneizziamo rispetto a t e otteniamo che la classificazione di $\tilde{f}$ come forma quadratica coincide con la classificazione di f come quadrica.

10.5 Esercizi

Esercizio 10.5.1. Mostrare che $\varphi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, con

$$\varphi((x_1, y_1, z_1), (x_2, y_2, z_2)) = x_1x_2 + 2x_1y_2 - z_1(y_2 + z_2)$$

è una forma bilineare e scriverne la matrice associata rispetto alle basi canoniche.

Esercizio 10.5.2. Sia

$$\varphi((x_1, y_1, z_1), (x_2, y_2, z_2)) = x_1x_2 + 2x_1y_2 + 2x_2y_1 - (y_1 + z_1)(y_2 + z_2).$$

Stabilire se è una forma bilineare simmetrica, dire se il vettore $(1, 1, 1)$ è isotropo per φ.

Esercizio 10.5.3. Sia

$$\varphi((x_1, y_1, z_1), (x_2, y_2, z_2)) = x_1x_2 + 2x_1y_2 + 2x_2y_1 - y_1y_2.$$

Stabilire se è una forma bilineare simmetrica, calcolare tutti i vettori isotropi di φ e stabilire se si tratta di una forma degenere.

Esercizio 10.5.4. La seguente forma quadratica è un prodotto scalare?

$$\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

$$\varphi((x_1, x_2), (y_1, y_2)) = (x_1 y_2 + x_2 y_1).$$

Calcolarne la segnatura e stabilire la base B rispetto alla quale $M_B(\varphi)$ è nella forma di Sylvester.

Esercizio 10.5.5. Sia $Q : \mathbb{R}^3 \to \mathbb{R}$ la seguente forma quadratica
$Q(x, y, z) = x^2 - y^2 + xz - 2yz$.
Dire se è definita positiva, negativa o indefinita e scriverla in forma canonica.

Esercizio 10.5.6. Qual è la forma normale della seguente forma quadratica? $Q : \mathbb{R}^4 \to \mathbb{R}$, $Q(x, y, z, t) = xy - yz + zt$. Qual è la base rispetto alla quale si può scrivere in forma normale?

Esercizio 10.5.7. Scrivere una base ortonormale rispetto alla seguente forma bilineare: $\varphi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, con
$\varphi((x_1, y_1, z_1), (x_2, y_2, z_2)) = (x_1 x_2 + 2y_1 y_2 - (1/2) z_1 z_2)$.

Esercizio 10.5.8. Classificare le seguenti coniche:

- $x^2 - xy - 3y^2 + 1 = 0$;
- $y^2 - xy + 3 = 0$;
- $x^2 + 2x - 4y + 5 = 0$;
- $2x^2 - xy + y - 1 = 0$;
- $x^2 + 2xy + y^2 = 0$.

Esercizio 10.5.9. Dire se le seguenti quadriche sono o no a centro:

- $xy - xz = 0$;
- $x^2 + xy - xz = 1$;
- $2x^2 + x - y + xz = 4$.

Esercizio 10.5.10. Classificare le seguenti quadriche:

- $x^2 + y^2 - 8z = 0$;
- $xy - z^2 = 9$;
- $4x + y - z^2 + x^2 = 1$;
- $x^2 - xz + 3z^2 = 0$;
- $xy - xz = 1$.

Esercizio 10.5.11. * Considerare la quadrica definita da:

$$\begin{pmatrix} x & y & z & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}.$$

Dire se tale quadrica sia o no degenere e classificarla.

CAPITOLO 11

Spazi affini

Prerequisiti: Sistemi lineari; spazi vettoriali; applicazioni lineari

Obiettivi: Saper utilizzare la generalizzazione di piano e spazio della ordinaria geometria analitica (come vista nel Capitolo 9) per spazi lineari a più dimensioni, tramite lo strumento degli spazi affini; saper passare dalla struttura di $\mathbb{R}^n$ come spazio vettoriale a quella di spazio affine $\mathbb{A}^n$.

11.1 Gli spazi affini

Facendo geometria analitica (nel piano, nello spazio), una parte essenziale è lo studio di punti, rette e piani (vedi Capitolo 9). Questi enti geometrici non sono sottospazi vettoriali di $\mathcal{V}_O^n$, o di $\mathbb{R}^n$, $n = 1, 2, 3$, tranne nel caso in cui essi contengano il vettore nullo. La struttura che generalizza ciò che si è visto in geometria analitica e che permette di trattare *sottospazi* diversi da quelli vettoriali è quella di *spazio affine*, che tratteremo brevemente in questo capitolo.

Definizione 11.1.1. Dato uno spazio vettoriale V, si dice *spazio affine su* V una tripla $(\mathbb{A}, V, \delta)$, dove $\mathbb{A}$ è un insieme e $\delta : \mathbb{A} \times \mathbb{A} \to V$ è una funzione (per indicare $\delta(P, Q)$ useremo talvolta $\overrightarrow{PQ}$, detto *vettore di punto iniziale* P e *punto finale* Q). La funzione δ è dotata delle proprietà seguenti:

1) $\forall P \in \mathbb{A}$, $\forall \mathbf{v} \in V$, esiste un unico $Q \in \mathbb{A}$ tale che $\mathbf{v} = \overrightarrow{PQ}$.

2) comunque dati $P, Q, R \in \mathbb{A}$, si ha $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$.

Useremo semplicemente $\mathbb{A}$ per indicare lo spazio affine quando V e δ siano sottointesi. La *dimensione* di $\mathbb{A}$, $\dim \mathbb{A}$, è definita come pari alla dimensione di V come spazio vettoriale.

Il prototipo del concetto di spazio affine è quello dei vettori geometrici.

Esempio 11.1.1. Se prendiamo $V = \mathcal{V}_O^3$, e $\mathbb{A}$ come l'insieme dei punti dello spazio della geometria euclidea, possiamo considerare ogni coppia di punti $(P, Q) \in \mathbb{A} \times \mathbb{A}$ e $\overrightarrow{PQ}$ sarà il vettore $\overrightarrow{OP'} \in \mathcal{V}_O^3$, uscente da O ottenuto traslando $\overrightarrow{PQ}$ nell'origine. In questo modo si ha che lavorare nello spazio affine $(\mathbb{A}, \mathcal{V}_O^3, \delta)$ corrisponde a considerare i punti dello spazio e i vettori $\overrightarrow{PQ}$); notiamo che per la proprietà 2), se $P = Q = R$, avremo che: $2\overrightarrow{PP} = \overrightarrow{PP}$, e quindi che $\forall P$, $\overrightarrow{PP}$ è il vettore nullo di $\mathcal{V}_O^3$, mentre per $R = P$, si avrà che $\overrightarrow{PQ} = -\overrightarrow{QP}$.

Esempio 11.1.2. Notiamo che ogni insieme $\mathbb{A} = \{P\}$, con un solo elemento, ha una struttura di spazio affine, con $\dim \mathbb{A} = 0$: $(\mathbb{A}, \{\mathbf{0}\}, \delta)$, ove δ è la funzione banale $\delta(P, P) = \mathbf{0}$. Gli spazi affini zerodimensionali sono detti *punti*.

Notiamo anche che per avere una struttura di spazio affine $(\mathbb{A}, V, \delta)$, l'insieme $\mathbb{A}$ deve poter essere messo in corrispondenza biunivoca con V. Infatti, sia $O \in \mathbb{A}$ un qualsiasi elemento; allora $\forall P \in \mathbb{A}$, avremo che esiste un unico $\mathbf{v} \in V$ tale che $\delta(O, P) = \mathbf{v}$; quindi la funzione $\delta(O, -) : \mathbb{A} \to V$ risulta una biezione fra $\mathbb{A}$ e V (che fa corrispondere il punto O a $\delta(O, O) = \overrightarrow{OO} = \mathbf{0} \in V$).

La Definizione 11.1.1 è molto generale (ad esempio potremmo avere diverse strutture di spazio affine sullo stesso spazio vettoriale o diverse strutture di spazio affine sullo stesso insieme $\mathbb{A}$); il seguente esempio mostra invece la struttura di spazio affine *più naturale* da considerare su V, data da $\mathbb{A} = V$.

Definizione 11.1.2. Dato uno spazio vettoriale V, consideriamo lo spazio affine su V: $\mathbb{A}(V)$, definito prendendo $\mathbb{A} = V$ e $\delta(\mathbf{v}, \mathbf{w}) = \mathbf{w} - \mathbf{v}$. La verifica che gli assiomi di spazio affine sono soddisfatti è immediata:

1) dati $\mathbf{p}, \mathbf{v} \in V$, l'elemento $\mathbf{q}$ cercato è: $\mathbf{q} = \mathbf{p} + \mathbf{v}$, infatti
$$\delta(\mathbf{p}, \mathbf{q}) = (\mathbf{p} + \mathbf{v}) - \mathbf{v} = \mathbf{p}.$$
2) $\forall \mathbf{p}, \mathbf{q}, \mathbf{r} \in V$, $\delta(\mathbf{p}, \mathbf{q}) + \delta(\mathbf{q}, \mathbf{r}) = \mathbf{q} - \mathbf{p} + \mathbf{r} - \mathbf{q} = \mathbf{r} - \mathbf{p} = \delta(\mathbf{p}, \mathbf{r})$.

Anche negli spazi affini useremo dei sistemi di coordinate, ricavate da una base dello spazio vettoriale associato.

Definizione 11.1.3. Sia $\mathbb{A}$ uno spazio affine su V, e sia $O \in \mathbb{A}$ e $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base di V. Diremo che $O, \mathbf{v}_1, \dots, \mathbf{v}_n$ è un *riferimento affine* di $\mathbb{A}$, e chiameremo *coordinate affini* di un punto $P \in \mathbb{A}$ le coordinate $a_1, \dots, a_n \in \mathbb{R}$ tali che $\overrightarrow{OP} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \in V$.

Tra tutti gli spazi affini, quelli che considereremo maggiormente sono i seguenti.

Definizione 11.1.4. Denoteremo con *spazi affini standard* gli spazi affini $\mathbb{A}^n = \mathbb{A}(\mathbb{R}^n)$; avremo che $\forall P, Q \in \mathbb{A}^n$, con $P = (a_1, \dots, a_n)$, $Q = (b_1, \dots, b_n)$, il vettore di punto iniziale P e punto finale Q è $\overrightarrow{PQ} = (b_1 - a_1, \dots, b_n - a_n) \in \mathbb{R}^n$.

Nello spazio affine $\mathbb{A}^n$ si dice *riferimento affine standard* il sistema di riferimento $O, e_1, \dots, e_n$, dato da $O = (0, \dots, 0) \in \mathbb{A}^n$ e dalla base canonica $\{e_1, \dots, e_n\}$ di $\mathbb{R}^n$.

Notiamo che in $\mathbb{A}^n$, rispetto al sistema di riferimento standard, ogni punto $P = (x_1, \dots, x_n)$ ha se stesso come n-pla di coordinate affini.

11.2 Sottospazi affini

Anche negli spazi affini abbiamo la nozione di sottospazio:

Definizione 11.2.1. Siano dati uno spazio affine $(\mathbb{A}, V, \delta)$, un sottospazio vettoriale $W \subset V$ e un punto $Q \in \mathbb{A}$; si dirà *sottospazio affine* di $\mathbb{A}$ *passante per* Q *e parallelo a* W, il sottoinsieme $S \subset \mathbb{A}$ tale che $S = \{P \in \mathbb{A} \mid \overrightarrow{PQ} \in W\}$. Porremo $\dim S = \dim W$, mentre la differenza $(\dim V - \dim W)$ è detta la *codimensione* del sottospazio.

Ovviamente si avrà che $Q \in S$, in quanto il vettore $\overrightarrow{QQ} = \mathbf{0} \in W$. Inoltre in $\mathbb{A}(V)$ il sottospazio affine passante per Q e parallelo a W si potrà scrivere: $S = \{Q + W\} = \{P \in \mathbb{A} \mid \exists \mathbf{w} \in W, P = Q + \mathbf{w}\}$, inoltre $\forall Q' \in S$, potremo anche scrivere $S = \{Q' + W\}$ (in quanto $\overrightarrow{QQ'} = Q - Q' \in W$, quindi $P = Q + \mathbf{w} = Q' + \mathbf{w} + (Q - Q') \in \{Q' + W\}$).

Avremo anche che S è uno spazio affine su W, considerando la tripla (S, W, δ'), ove δ' è la restrizione di δ a $S \times S$.

Esempio 11.2.1. Consideriamo i sottospazi $W_1, W_2 \subset \mathbb{R}^3$, definiti da

$$W_1 = \{(x,y,z) \,|\, x+y+z=0\}, W_2 = \{(x,y,z) \,|\, x+y=0, y+z=0\}.$$

Sia $Q = (1,2,3) \in \mathbb{A}^3$, e sia $S_i \subset \mathbb{A}$ il sottospazio affine per Q parallelo a W_i, $i = 1,2$, definito da: $S_i = \{P \in \mathbb{A}^3 \,|\, \overrightarrow{PQ} \in W_i\}$, ove si avrà $\overrightarrow{PQ} = (x-1, y-2, z-3)$; allora:

$$S_1 = \{(x,y,z) \in \mathbb{A}^3 \,|\, x+y+z=6\};$$

$$S_2 = \{(x,y,z) \in \mathbb{A}^3 \,|\, x+y=3, y+z=5\}.$$

In questo caso i sottospazi vettoriali W_1, W_2 rappresentano, rispettivamente, un piano e una retta (per l'origine) nello spazio tridimensionale; i sottospazi affini S_1, S_2 in $\mathbb{A}^3$ rappresentano il piano e la retta paralleli ai precedenti e passanti per Q (vedi anche Capitolo 9).

Generalizziamo quanto visto nell'esempio precedente: sia $W \subset \mathbb{R}^n$ un sottospazio vettoriale definito tramite un sistema lineare omogeneo $AX = O$, ove $A \in \mathbb{R}^{m,n}$, $X = {}^t(x_1, \ldots, x_n)$ e $O = {}^t(0, \ldots, 0) \in \mathbb{R}^{m,1}$; cioè W sia costituito dalle soluzioni del sistema $AX = O$. Allora, dato $Q = (a_1, \ldots, a_n) \in \mathbb{A}^n$, il sottospazio affine $S \subset \mathbb{A}^n$ passante per Q e parallelo a W, sarà definito dall'insieme:

$$S = \{P = (x_1, \ldots, x_n) \,|\, \overrightarrow{PQ} \in W\} = \{(x_1, \ldots, x_n) \,|\, A(X-Q) = \mathbf{0}\} =$$

$$= \{(x_1, \ldots, x_n) \,|\, AX = B\},$$

ove $B = A \cdot {}^t(a_1, \ldots, a_n) \in \mathbb{R}^{m,1}$.

Il sottospazio S verrà detto una *retta affine* o un *piano affine* se, rispettivamente, $\dim S = \dim W = 1$ o 2; verrà invece detto un *iperpiano affine* se $\dim S = n-1$.

Dato che ogni sottospazio vettoriale $W \subset \mathbb{R}^n$ si può definire tramite un sistema omogeneo $AX = O$, $A \in \mathbb{R}^{m,n}$, quanto appena visto sopra ci dice che ogni sottospazio affine $S \subset \mathbb{A}^n$ si può definire tramite un sistema lineare $AX = B$.

In particolare, in quanto $\dim W = n - r$, ove r sia il rango della matrice A che definisce W come spazio delle soluzioni di $AX = O$; avremo che gli iperpiani di $\mathbb{A}^n$ si definiscono con un'unica equazione $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$, ove $(a_1, \ldots, a_n) \neq O$.

Esempio 11.2.2. Se consideriamo in $\mathbb{A}^5$ (con il riferimento standard) le equazioni $x_1 + x_2 + x_3 = 3$ e $x_4 + x_5 = 2$, esse definiranno due iperpiani affini S_1 e S_2, ciascuno di dimensione 4. Avremo che $S = S_1 \cap S_2$ sarà definito dal sistema formato dalle due equazioni, e sarà un sottospazio affine di dimensione 3.

Vediamo inoltre che intersecando sottospazi affini otterremo ancora sottospazi affini.

Proposizione 11.2.1. *Sia $\mathbb{A}(V)$ lo spazio affine su V (vedi 11.1.2), siano $W_1, \ldots, W_t$ sottospazi vettoriali di V e siano S_i, $i = 1, \ldots, t$ i sottospazi affini di $\mathbb{A}(V)$, paralleli a W_i e passanti, rispettivamente, per $Q_i \in \mathbb{A}$. Allora l'intersezione $\bigcap_{i=1}^{t} S_i$, se non è vuota, è un sottospazio affine di $\mathbb{A}$.*

Dimostrazione. Possiamo scrivere $S_i = \{Q_i + W_i\}$; allora se $S = \bigcap_{i=1}^{t} S_i \neq \emptyset$, sia $Q \in S$; per quanto visto prima possiamo anche scrivere $S_i = \{Q + W_i\}$, quindi

$$\bigcap_{i=1}^{t} S_i = \bigcap_{i=1}^{t} \{Q + W_i\} = Q + \bigcap_{i=1}^{t} W_i,$$

che è un sottospazio affine (l'intersezione di sottospazi vettoriali è un sottospazio vettoriale). □

Possiamo ottenere sottospazi affini anche partendo da punti, vediamolo in $\mathbb{A}^n$.

Definizione 11.2.2. Siano $P_0, P_1, \ldots, P_s \in \mathbb{A}^n$, $s \geq 1$. Il *sottospazio affine generato da* $P_0, \ldots, P_n$ è il sottospazio affine passante per P_0 e parallelo a $\langle \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \ldots, \overrightarrow{P_0P_s} \rangle \subset \mathbb{R}^n$. Si indica talvolta con $\overline{P_0, \ldots, P_s}$.

Esempio 11.2.3. Consideriamo i punti

$P_0 = (0, 1, 2), P_1 = (-1, 2.1), P_2 = (1, 0, 3)$,

e lo spazio $\overline{P_0, P_1, P_2} \subset \mathbb{A}^3$. Avremo che tale sottospazio è parallelo a $W = \langle (1, -1, 1), (-1, 1, -1) \rangle = \langle (1, -1, 1) \rangle$, quindi si ha

$\dim \overline{P_0, P_1, P_2} = 1$, cioè $\overline{P_0, P_1, P_2}$ è una retta nello spazio affine.

I tre punti risultano perciò allineati.

In generale, avremo che $\dim \langle \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \ldots, \overrightarrow{P_0P_s} \rangle \leq s$, quindi anche $\dim \overline{P_0, \ldots, P_s} \leq s$.

Definizione 11.2.3. Siano $P_0, P_1, \ldots, P_s \in \mathbb{A}^n$, quando vale l'uguaglianza $\dim \overline{P_0, \ldots, P_s} = s$, $P_0, \ldots, P_s$ si dicono *indipendenti* (ciò implica che $s \leq n$).

Nell'esempio precedente abbiamo visto tre punti non indipendenti. Dalla definizione è immediato ricavare che $P_0, \ldots, P_s$ sono indipendenti in $\mathbb{A}^n$ se e solo se $\overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \ldots, \overrightarrow{P_0P_s}$ sono linearmente indipendenti in $\mathbb{R}^n$ (ciò equivale al fatto che $\dim\langle \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, \ldots, \overrightarrow{P_0P_s}\rangle = s$). Naturalmente se $s > n$, ciò non può accadere.

Esempio 11.2.4. Siano $P_0 = (0,1,2,0), P_1 = (-1,2.1,1), P_2 = (1,0,3,1)$ in $\mathbb{A}^4$; avremo che $\overline{P_0,P_1,P_2}$ è il sottospazio affine per P_0 parallelo a $W = \langle(-1,1,-1,1),(1,-1,1,1)\rangle$; si ha che $\dim W = 2$ (i due generatori sono linearmente indipendenti). Quindi $\overline{P_0,P_1,P_2}$ è un piano affine in $\mathbb{A}^4$, e i tre punti sono indipendenti.

Un altro modo per individuare sottospazi affini è mediante loro equazioni parametriche.

Definizione 11.2.4. Dato il sottospazio affine $S \subset \mathbb{A}^n$, parallelo a $W \subset \mathbb{R}^n$ e passante per $Q = (q_1, \ldots, q_n) \in \mathbb{A}^n$ (usando il riferimento standard), sia $\{\mathbf{w}_1, \ldots, \mathbf{w}_s\}$ una base di W, con $\mathbf{w}_i = (w_{1i}, \ldots, w_{ni})$. Per ogni $P \in S$, esisteranno $t_1, \ldots, t_s \in \mathbb{R}$ tali che:

$$\overrightarrow{QP} = t_1 w_1 + \cdots + t_s w_s$$

poiché $\overrightarrow{PQ} = (x_1 - q_1, \ldots, x_n - q_n)$, avremo:

$$\begin{cases} x_1 = t_1 w_{11} + \cdots + t_n w_{1s} + q_1 \\ x_2 = t_1 w_{21} + \cdots + t_n w_{2s} + q_2 \\ \qquad\qquad \cdots \\ x_n = t_1 w_{n1} + \cdots + t_n w_{ns} + q_n \end{cases}$$

Queste sono dette *equazioni parametriche di* S. Ogni $(t_1, \ldots, t_s) \in \mathbb{R}^s$ dà un punto $(x_1, \ldots, x_n) \in S$.

Ovviamente le equazioni parametriche di S non sono uniche, ma dipendono dalla scelta di Q e della base di W.

Esempio 11.2.5. Sia $S \subset \mathbb{A}^4$ il sottospazio passante per $Q = (1,2,3,4)$ e parallelo a $W = \{(x,y,z,t) \in \mathbb{R}^4 | x + y = 0, x + z - t = 0\}$. Una base di W è data da $\{(-1,1,1,0),(1,-1,0,1)\}$ ($\dim S = \dim W = 2$); sono equazioni parametriche per S:

$$\begin{cases} x = -t_1 + t_2 + 1 \\ y = t_1 - t_2 + 2 \\ z = t_1 + 3 \\ t = t_2 + 4 \end{cases}$$

11.3 Affinità

Consideriamo adesso le applicazioni fra spazi affini.

Definizione 11.3.1. Dati due spazi affini $(\mathbb{A}, V, \delta)$, $(\mathbb{A}', V', \delta')$, un *isomorfismo* fra $\mathbb{A}$ e $\mathbb{A}'$ è un'applicazione biunivoca $f : \mathbb{A} \to \mathbb{A}'$ tale che esista un isomorfismo $\varphi : V \to V'$ di spazi vettoriali per cui ($\forall P, Q \in \mathbb{A}$):

$$\delta'(f(P), f(Q)) = \varphi(\delta(P, Q)), \qquad \text{i. e.} \qquad \overrightarrow{f(P)f(Q)} = \varphi(\overrightarrow{PQ}).$$

Due spazi affini fra cui esiste un isomorfismo si dicono *isomorfi*. Chiameremo poi *un'affinità* di $\mathbb{A}$ un isomorfismo di $\mathbb{A}$ su se stesso.

Non è difficile vedere che l'identità $I_{\mathbb{A}} : \mathbb{A} \to \mathbb{A}$ è un'affinità, che se $f : \mathbb{A} \to \mathbb{A}'$ è un'affinità lo è anche $f^{-1} : \mathbb{A}' \to \mathbb{A}$, e se $g : \mathbb{A}' \to \mathbb{A}''$ è un'altra affinità, anche $g \circ f : \mathbb{A} \to \mathbb{A}''$ lo è .

Si può notare che se $\mathbb{A}$ è uno spazio affine, con $\dim \mathbb{A} = n$, allora, scelto un sistema di riferimento affine $O, \mathbf{v}_1, \ldots, \mathbf{v}_n$ di $\mathbb{A}$, l'applicazione $f : \mathbb{A} \to \mathbb{A}^n$ che associa ad ogni punto $P \in \mathbb{A}$ la sua n-upla di coordinate è un isomorfismo di spazi affini (l'isomorfismo di spazi vettoriali associato è $\varphi : V \to \mathbb{R}^n$, ove $\forall v = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n \in V$, $\varphi(\mathbf{v}) = (x_1, \ldots, x_n)$.

Si ha quindi che ogni spazio affine di dimensione n è isomorfo ad $\mathbb{A}^n$ e quindi che ogni volta che dobbiamo lavorare in uno spazio affine di dimensione n potremo, fissato un suo sistema di riferimento, trasferirci a lavorare su $\mathbb{A}^n$, a meno di isomorfismi.

Fra tutte le affinità che possiamo considerare su $\mathbb{A}^n$, (come su ogni altro spazio affine), sono rilevanti quelle dette traslazioni:

Definizione 11.3.2. Sia $\mathbf{v} \in \mathbb{R}^n$; l'affinità $t_{\mathbf{v}}$ di $\mathbb{A}^n$, che associa ad ogni $P \in \mathbb{A}^n$ il punto $t_{\mathbf{v}}(P)$ tale che $\overrightarrow{Pt_{\mathbf{v}}(P)} = \mathbf{v}$, allora $t_{\mathbf{v}}$ si dice *traslazione definita da* $\mathbf{v}$.

Possiamo verificare che $t_{\mathbf{v}} : \mathbb{A}^n \to \mathbb{A}^n$ è un'affinità: come prima cosa verifichiamo che $t_{\mathbf{v}}$ è biunivoca. Per ogni $Q \in \mathbb{A}^n$, sia $P = t_{-\mathbf{v}}(Q)$, avremo allora che $Q = t_{\mathbf{v}}(P)$ in quanto $\overrightarrow{QP} = -\mathbf{v}$, quindi $\overrightarrow{PQ} = \mathbf{v}$. Allora $t_{-\mathbf{v}} = t_{\mathbf{v}}^{-1}$ e $t_{\mathbf{v}}$ è biunivoca.

Abbiamo poi:

$$\overrightarrow{t_{\mathbf{v}}(P)t_{\mathbf{v}}(Q)} = \overrightarrow{t_{\mathbf{v}}(P)P} + \overrightarrow{PQ} + \overrightarrow{Qt_{\mathbf{v}}(Q)} = -\mathbf{v} + \overrightarrow{PQ} + \mathbf{v} = \overrightarrow{PQ} :$$

quindi l'isomorfismo di $\mathbb{R}^n$ associato a $t_{\mathbf{v}}$ è l'identità e $t_{\mathbf{v}}$ è un'affinità (in particolare $t_{\mathbf{0}}$ è l'identità).

Notiamo anche che se $f : \mathbb{A}^n \to \mathbb{A}^n$ è un'affinità associata all'identità $id_{\mathbb{R}^n}$, allora $\forall P, Q \in \mathbb{A}^n$, avremo $\overrightarrow{f(P)f(Q)} = \overrightarrow{PQ}$, il che implica anche che il vettore $\mathbf{v} = \overrightarrow{Pf(P)} = \overrightarrow{Qf(Q)}$ sia indipendente dalla scelta di P e di Q, e quindi $f = t_{\mathbf{v}}$.

Abbiamo in questo modo dimostrato che le traslazioni sono esattamente le affinità di $\mathbb{A}^n$ che sono associate all'identità $id_{\mathbb{R}^n}$.

Vediamo che anche per le affinità di $\mathbb{A}^n$ si può dare una descrizione matriciale. Sia $f : \mathbb{A}^n \to \mathbb{A}^n$ un'affinità e $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ l'isomorfismo di spazi vettoriali ad essa associato. Usiamo la base canonica $\mathcal{E}$ su $\mathbb{R}^n$ e l'associato riferimento standard su $\mathbb{A}^n$; sia $A = M_{\mathcal{E}\mathcal{E}}(f) \in \mathbb{R}^{n,n}$ la matrice associata a φ, e sia $f(O) = B \in \mathbb{A}^n$. Allora $\forall X \in \mathbb{A}^n$ si ha:

$$f(X) - B = f(X) - f(O) = A(X - O) = AX$$

quindi l'affinità si può scrivere (rispetto al riferimento affine $Oe_1, \ldots, e_n$):

$$f(X) = AX + B.$$

Notiamo che se f è una traslazione $t_{\mathbf{v}}$, allora $A = I_n$ (la matrice identità in $\mathbb{R}^{n,n}$) e quindi avremo

$$t_{\mathbf{v}}(X) = X + B$$

ove $B = t_{\mathbf{v}}(O)$, quindi $\overrightarrow{OB} = \mathbf{v}$; inoltre, in quanto stiamo usando il riferimento affine standard su $\mathbb{A}^n$ e la base canonica in $\mathbb{R}^n$, abbiamo che B e $\mathbf{v}$ si rappresentano con la stessa n-upla $(b_1, \ldots, b_n)$.

Poiché ogni affinità si può scrivere come $f(X) = AX + B$, ciò si può pensare come: ogni affinità su $\mathbb{A}^n$ si ottiene (scritta rispetto al sistema di riferimento standard) dalla applicazione lineare $X \to AX$, composta con la traslazione $t_{\overrightarrow{OB}}$. Ciò ci dice che quando consideriamo le affinità come funzioni sull'insieme $\mathbb{A}^n = \mathbb{R}^n$, stiamo considerando le composizioni di isomorfismi lineari e traslazioni (ogni affinità si compone di due di essi, come appena notato).

11.4 Esercizi

Esercizio 11.4.1. Determinare la dimensione del sottospazio affine di $\mathbb{A}^4$ determinato dalle soluzioni del sistema:

$$\begin{cases} 3x + 2y - z - t = 1 \\ x + y + z + t = 2 \\ 2x + y - 2z - 2t = -1 \end{cases}$$

Esercizio 11.4.2. Dare un esempio di tre sottospazi affini di $\mathbb{A}^3$, diversi da tutto lo spazio e aventi dimensioni diverse fra loro.

Esercizio 11.4.3. Esprimere (nel riferimento standard) il sottospazio affine $S \subset \mathbb{A}^4$ dell'esempio 11.2.5 tramite un sistema del tipo $AX = B$

Esercizio 11.4.4. Determinare (con equazioni parametriche o come soluzioni di un sistema) il sottospazio affine $S_1 = \overline{P_0, P_1} \subset \mathbb{A}^4$, ove siano $P_0 = (0,1,1,0)$, $P_1 = (1,0,0,1)$. Determinare anche l'intersezione di S_1 con l'iperpiano $S_2 \subset \mathbb{A}^4$ definito dall'equazione (nel riferimento standard): $x_1 + x_2 + x_3 - x_4 = 2$.

Esercizio 11.4.5. Dimostrare che un sottospazio affine S di $\mathbb{A}^n$ è anche un sottospazio vettoriale di $\mathbb{R}^n$ se e solo se $\mathbf{0} \in S$.

Esercizio 11.4.6. Sia W un sottospazio vettoriale di $\mathbb{R}^n$. W possiede anche la struttura di sottospazio affine di $\mathbb{A}^n$?

Esercizio 11.4.7. $\mathbb{A}^6$ è un sottospazio affine di $\mathbb{A}^7$?

Esercizio 11.4.8. * Siano $P_0, \ldots, P_s \in \mathbb{A}^n$; considerare i sottospazi di $\mathbb{A}^n$: $S_1 = \overline{P_0, \ldots, P_s}$ e $S_2 = \overline{P_0, \ldots, P_{s-1}}$. Dire se le seguenti affermazioni siano vere o false (motivando la risposta).

1) S_2 è un sottospazio affine di S_1;
2) $\dim S_2 = \dim S_1 - 1$;
3) $\dim S_1 = s$.

Esercizio 11.4.9. Sia S un sottospazio affine di $\mathbb{A}^n$ definito (nel riferimento standard) dal sistema $AX = B$. Dimostrare che codimensione di S è pari al rango di A.

Esercizio 11.4.10. Abbiamo visto che se $f : \mathbb{A}^n \to \mathbb{A}^n$ è un'affinità associata all'identità $id_{\mathbb{R}^n}$, allora $\forall P, Q \in \mathbb{A}^n$, avremo $\overrightarrow{f(P)f(Q)} = \overrightarrow{PQ}$. Si dimostri che ciò implica anche che il vettore $\mathbf{v} = \overrightarrow{Pf(P)} = \overrightarrow{Qf(Q)}$ sia indipendente dalla scelta di P e di Q (e quindi $f = t_{\mathbf{v}}$).

Esercizio 11.4.11. Sia V un sottospazio vettoriale di $\mathbb{R}^n$; esso ha anche una struttura di sottospazio affine di $\mathbb{A}^n$?

CAPITOLO 12

Numeri complessi

Prerequisiti: Nessun prerequisito.

Obiettivi: Saper utilizzare i numeri complessi nelle loro varie espressioni (algebrica, cartesiana, trigonometrica, esponenziale); saper risolvere semplici equazioni a valori complessi.

12.1 Introduzione

In questo libro abbiamo sempre lavorato sui numeri reali. Questo capitolo è invece interamente dedicato allo studio dei numeri complessi.

Iniziamo con un semplice esempio.

Consideriamo una qualunque equazione di secondo grado a valori reali in una variabile x:

$$ax^2 + bx + c = 0. \tag{12.1}$$

Sappiamo che una tale equazione non sempre ammette soluzioni (reali); essa ha soluzioni solo se $b^2 - 4ac \geq 0$. Vogliamo allora ampliare il campo dei numeri reali in modo da ottenere un nuovo campo, $\mathbb{C}$, dove le equazioni di secondo grado come (12.1) abbiano *sempre* soluzioni.

Partiamo dal caso particolare in cui nella (12.1) si abbia $b = 0$, mentre $a = c = 1$, cosicché l'equazione sia $x^2 + 1 = 0$; tale equazione non ha soluzione in $\mathbb{R}$. Affinché l'avesse dovrebbe esistere un numero z tale che $z^2 = -1$, ma, come sappiamo, ciò è impossibile in $\mathbb{R}$. Introdurremo allora un "nuovo numero": i (quella che in genere si chiama *unità immaginaria*), che aggiungeremo all'insieme $\mathbb{R}$, e la cui principale proprietà sarà quella di avere il quadrato uguale a -1:

$$i^2 = -1.$$

Storicamente, i numeri complessi nascono proprio dall'esigenza di saper

trattare radici di numeri negativi; quest'esigenza si presentò nel XVI secolo con la scoperta della formula risolutiva per le equazioni di terzo grado (Scipione del Ferro, Nicolò Tartaglia, Gerolamo Cardano); tale formula contempla delle radici sotto le quali possono apparire quantità negative, come accade in quelle di secondo grado. In questo caso può capitare però che, dopo opportuni passaggi, tali radici si possano semplificare portando a un accettabile risultato reale: accade quindi che per arrivare alle soluzioni (reali) dell'equazione bisogna operare con radici di numeri negativi, e da qui la necessità di introdurre i numeri complessi. Per lungo tempo essi vennero considerati solo degli artifici algebrici per la risoluzione delle equazioni, e solo nell'800, grazie alla sistemazione fattane da C. Gauss, furono considerati dei numeri veri e propri, alla stessa stregua dei reali.

12.2 Rappresentazione algebrica

Quanto visto sopra è il primo passo nella costruzione dell'insieme $\mathbb{C}$; l'idea è quella di partire con l'aggiungere a $\mathbb{R}$ un nuovo elemento i, la cui caratteristica sia che, con l'operazione di prodotto che definiremo nel nuovo insieme, $i^2 = -1$.

Dal momento che vogliamo che nel nuovo insieme si estendano le operazioni di somma e prodotto già presenti in $\mathbb{R}$, dovremo avere degli elementi che corrispondono a tutti i prodotti fra i ed ogni numero reale b, quindi in $\mathbb{C}$ inseriremo tutti i simboli ib, $\forall\, b \in \mathbb{R}$, detti numeri *immaginari*. Dovremo avere anche tutte le possibili somme fra un numero reale a ed un ib; definiamo allora l'insieme $\mathbb{C}$ come dato da tutte le possibili espressioni del tipo $a + ib$, ove a, b siano numeri reali; intenderemo che $a = a + 0i$ e che $ib = 0 + ib$.

Si userà in genere la variabile z per rappresentare un generico elemento di $\mathbb{C}$ (come si usa la x in $\mathbb{R}$); scriveremo quindi $z = a + ib$ per ogni $z \in \mathbb{C}$. Quanto ottenuto è descritto dalla seguente definizione:

Definizione 12.2.1. L' insieme:

$$\mathbb{C} = \{z = a + ib \,|\, a, b \in \mathbb{R}\},$$

è detto *insieme dei numeri complessi*. In particolare, a viene chiamata *parte reale* di z ed indicata con $Re(z)$, mentre b viene chiamata *parte immaginaria* di z ed indicata con $Im(z)$. Il numero complesso $i \in \mathbb{C}$ è chiamato *unità immaginaria*.

Con questa definizione è individuato l'*insieme* $\mathbb{C}$; notiamo che il sottoinsieme di $\mathbb{C}$ dato dagli z tali che $Im(z) = 0$ coincide con l'insieme $\mathbb{R}$, che risulta quindi un sottoinsieme di $\mathbb{C}$.

La struttura algebrica $(\mathbb{C}, +, \cdot)$ è data definendo le operazioni di somma e prodotto su $\mathbb{C}$. Questo si fa adottando le stesse regole di calcolo che si usano in $\mathbb{R}[x]$ (cioè trattando ogni $z = a + ib$ come un polinomio in i) ma sostituendo i^2 con -1 quando appare):

- $i \cdot i = i^2 = -1$;
- $\forall\, z = a + ib, w = c + id \in \mathbb{C}$, si ha:
  $z + w = (a + c) + i(b + d)$;
  $z \cdot w = (a + ib)(c + id) = ac + i(ad + bc) + i^2 bd =$
  $= (ac - bd) + i(ad + bc)$.

È facile notare che le operazioni di somma e prodotto così definite coincidono con quelle usuali in $\mathbb{R}$ quando effettuate su numeri complessi aventi $Im(z) = 0$, e che quindi $(\mathbb{R}, +, \cdot)$ è una sottostruttura di $(\mathbb{C}, +, \cdot)$.

Vediamo ora che la struttura algebrica così definita è un *campo*. Le proprietà dell'addizione sono ovvie (e lasciamo la loro dimostrazione come semplice esercizio):

- Essa è commutativa e associativa;
- L'elemento neutro dell'addizione è l'elemento $0 = 0 + i0$;
- L'opposto di un elemento $a + ib \in \mathbb{C}$ è $-a - ib = -a + i(-b)$;

Queste proprietà rendono la struttura $(\mathbb{C}, +)$ un *gruppo commutativo*.

Passiamo ora alle proprietà della moltiplicazione (anch'esse lasciate come esercizio):

- Essa è commutativa, associativa e distributiva rispetto alla somma;
- L'elemento neutro della moltiplicazione è l'elemento $1 = 1 + i0$;
- L'inverso, rispetto al prodotto, di un elemento $z = a + ib \in \mathbb{C}$, $z \neq 0$, è
$$z^{-1} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2};$$
infatti, $\forall\, (a, b) \in \mathbb{R}^2$, $(a, b) \neq (0, 0)$, si ha che
$$(a + ib).\left(\frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}b\right) = \frac{a^2}{a^2 + b^2} - i\frac{ab}{a^2 + b^2} +$$
$$+i\frac{ab}{a^2 + b^2} - i^2\frac{b^2}{a^2 + b^2} = \frac{a^2}{a^2 + b^2} + \frac{b^2}{a^2 + b^2} = 1.$$

Con queste ulteriori proprietà abbiamo che la struttura $(\mathbb{C}, +, \cdot)$ è un *campo*.

Risulterà utile il concetto di coniugato di un numero complesso:

Definizione 12.2.2. Per ogni numero complesso $z = a + ib \in \mathbb{C}$ con $a, b \in \mathbb{R}$, si definisce il *coniugato* di z come l'elemento $a - ib \in \mathbb{C}$ e lo si indica con $\overline{z}$.

Elenchiamo qui alcune proprietà che non dimostreremo ma che lasciamo come facile esercizio.

Proposizione 12.2.1. *Per ogni $z = a+ib, z_1, z_2 \in \mathbb{C}$ valgono le seguenti proprietà :*

- $z + \overline{z} = 2a = 2Re(z)$;
- $z - \overline{z} = 2b = 2Im(z)$;
- $\overline{\overline{z}} = z$;
- $z.\overline{z} = a^2 + b^2 = (Re(z))^2 + (Im(z))^2$;
- $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$;
- $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$;
- $\overline{z^{-1}} = \overline{z}^{-1}$.

È inoltre possibile definire la nozione di *modulo* di un numero complesso. Abbiamo poc'anzi osservato che $z \cdot \overline{z} = (Re(z))^2 + (Im(z))^2$ che è ovviamente un numero reale sempre positivo per ogni $z \in \mathbb{C}$; ha quindi senso estrarne la radice quadrata anche nei numeri reali, e adotteremo tale radice come "modulo" per un numero complesso.

Definizione 12.2.3. Il *modulo* $|z| \in \mathbb{R}$, $|z| \geq 0$ di un numero complesso $z = a + ib \in \mathbb{C}$ è definito come segue:

$$|z| = \sqrt{z.\overline{z}} = \sqrt{a^2 + b^2} = \sqrt{(Re(z))^2 + (Im(z))^2} \in \mathbb{R}^+.$$

Osserviamo che se $Im(z) = 0$, ossia se $z = a$ è un numero reale, allora $|z| = \sqrt{a^2} = |a|$ che è la normale nozione di modulo nei numeri reali.

Vediamo alcune proprietà del modulo per i numeri complessi. Per ogni $z = a + ib, z_1 = a_1 + ib_1, z_2 = a_2 + ib_2 \in \mathbb{C}$ valgono le seguenti:

- $|z| \geq 0$;
- $|z| = 0$ se e solo se $z = 0$;
- $|Re(z)| \leq |z|$, $|Im(z)| \leq |z|$, $|z| \leq |Re(z)| + |Im(z)|$;
- $|z_1.z_2| = |z_1|.|z_2|$;
- $\left|z^{-1}\right| = |z|^{-1}$;
- $z^{-1} = \overline{z}|z|^{-1}$;
- $|z_1 + z_2| \leq |z_1| + |z_2|$ (disuguaglianza triangolare);
- $|z_1 + z_2| \geq ||z_1| - |z_2||$.

12.3 Rappresentazione cartesiana

I numeri complessi possono essere rappresentati in diversi modi tra loro tutti equivalenti. Iniziamo in questa sezione a trattare i numeri complessi in quella che viene chiamata la *rappresentazione cartesiana*: mentre i numeri reali si rappresentano in corrispondenza biunivoca con i punti di una retta, per rappresentare i complessi si utilizzano i punti di un piano.

Definizione 12.3.1. I numeri complessi $\mathbb{C}$ possono essere messi in corrispondenza biunivoca con l'insieme $\mathbb{R}^2 = \{(a, b) \,|\, a, b \in \mathbb{R}\}$; in tale corrispondenza $z = a + ib \to (a, b)$, e le operazioni di addizione e moltiplicazione che considereremo su $\mathbb{R}^2$ sono:

- Addizione: $(a, b) + (c, d) = (a + c, b + d)$;
- Moltiplicazione: $(a, b) \cdot (c, d) = (ac - bd, ad + bc)$.

Notiamo subito che l'operazione di somma è quella che normalmente si utilizza in $\mathbb{R}^2$ quando lo si considera come spazio vettoriale; quindi abbiamo già visto che

- essa è commutativa e associativa;
- l'elemento neutro dell'addizione è l'elemento $u = (0, 0)$ ossia $\forall\, (a, b) \in \mathbb{C}$ si ha che $(a, b) + u = u + (a, b) = (a, b)$;
- l'opposto di un elemento $(a, b) \in \mathbb{C}$ è $(-a, -b)$: $\forall\, (a, b) \in \mathbb{C}$ si ha che $(a, b) + (-a, -b) = (-a, -b) + (a, b) = u$.

La moltiplicazione che abbiamo definito risulta invece un'operazione del tutto nuova rispetto a quelle che abbiamo visto su $\mathbb{R}^2$ come spazio vettoriale (o anche vettoriale euclideo); ecco alcune proprietà della moltiplicazione (che lasciamo come semplice esercizio):

- essa è commutativa, associativa e distributiva rispetto alla somma;
- l'elemento neutro della moltiplicazione è l'elemento $(1,0)$: ossia $\forall\,(a,b) \in \mathbb{C}$ si ha che $(a,b)\cdot(1,0) = (1,0)\cdot(a,b) = (a,b)$;
- notiamo che se moltiplichiamo per un numero del tipo $(\lambda,0)$, allora si ha: $\forall\,(a,b),(\lambda,0) \in \mathbb{C}$, $(\lambda,0)\cdot(a,b) = (\lambda a, \lambda b)$, in questo caso scriveremo semplicemente $\lambda(a,b) = (\lambda a, \lambda b)$, identificando le coppie $(\lambda,0)$ con i numeri reali λ, e quindi anche $(1,0)$ con 1;
- l'inverso di un elemento $(a,b) \in \mathbb{C}$, $(a,b) \neq (0,0)$, è $\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$: $\forall\,(a,b) \in \mathbb{C}$ si ha che $(a,b)\cdot\left(\frac{a}{a^2+b^2}, \frac{b}{a^2+b^2}\right) = \left(\frac{a}{a^2+b^2}, \frac{b}{a^2+b^2}\right)\cdot(a,b) = \tilde{u}$.

Notiamo che con queste operazioni, $\mathbb{R}^2$ ha la struttura di *campo*, e la corrispondenza $z \in \mathbb{C} \to (a,b)$ con $\mathbb{R}^2$ conserva le rispettive operazioni; si ha quindi che $\mathbb{C}$ può venir rappresentato sul piano cartesiano, come $\mathbb{R}^2$, e ogni numero complesso $z = a+ib$ si visualizza geometricamente con un punto di coordinate (a,b) o con un vettore (tramite l'identificazione con $\mathcal{V}_O^2$, vedi Figura 12.1).

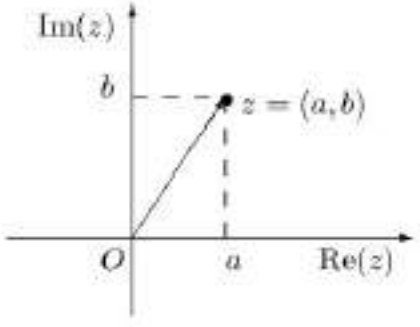


Figura 12.1: Rappresentazione cartesiana di un numero complesso z con $\mathrm{Re}(z) = a$ e $\mathrm{Im}(z) = b$.

Chiameremo l'asse delle x *asse reale* e l'asse delle y *asse immaginario* di $\mathbb{C}$, in quanto i loro punti rappresentano, rispettivamente, i numeri reali $(\lambda,0)$ e quelli *immaginari* $(0,b)$.

Definizione 12.3.2. La corrispondenza data nella Definizione 12.3.1 si chiama *rappresentazione cartesiana* dei numeri complessi.

Notiamo che il campo $\mathbb{C}$ dei numeri complessi ha anche la struttura di spazio vettoriale su $\mathbb{R}$ (isomorfa a quella di $\mathbb{R}^2$); come spazio vettoriale

si ha che $\dim \mathbb{C} = 2$, ed una base è data da $\{1, i\}$, o $\{(1,0),(0,1)\}$, la base canonica, nella rappresentazione cartesiana.

Osserviamo come con questa interpretazione, l'addizione tra due numeri complessi coincida con la somma tra vettori che ormai ben conosciamo (vedi Figura 12.2).

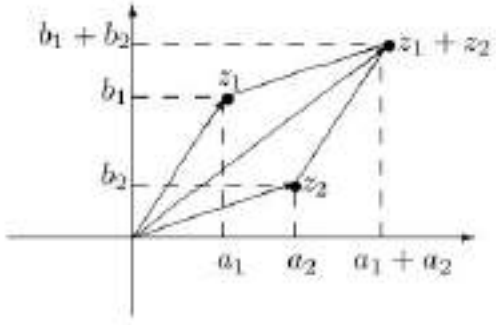


Figura 12.2: Somma dei numeri complessi $z_1 = (a_1, b_1)$ e $z_2 = (a_2, b_2)$.

Anche la moltiplicazione tra numeri complessi ha una rappresentazione significativa sul piano cartesiano, ma la vedremo più avanti.

12.4 Rappresentazione trigonometrica

Nelle rappresentazioni cartesiana e algebrica abbiamo descritto un numero complesso z o tramite le sue coordinate nel piano complesso (cioè $z = (a, b)$) o come somma della sua parte reale e della sua parte immaginaria $z = a + ib$. Vogliamo descrivere ora i numeri complessi tramite il loro modulo e l'angolo che staccano con l'asse delle x quando rappresentati nel piano complesso (in analogia con l'uso delle *coordiante polari* in $\mathbb{R}^2$).

Consideriamo l'insieme $\{w \in \mathbb{C} \,|\, |w| = 1\}$. Esso individua tutti i numeri complessi che giacciono sulla circonferenza unitaria centrata nell'origine degli assi cartesiani.

Chiamiamo θ l'angolo individuato dall'asse reale (l'asse dei punti di tipo $(a, 0)$ con $a \in \mathbb{R}$) e la retta passante per l'origine e w e orientiamo tale angolo in senso antiorario (vedi Figura 12.3). In questo modo è facile vedere che $Re(w) = \cos\theta$, $Im(w) = \sin\theta$, quindi il nostro vettore unitario w può essere anche scritto come $w = \cos\theta + i\sin\theta$.

La corrispondenza tra gli angoli θ nell'intervallo $[0, 2\pi[$ e gli elementi della circonferenza $\{w \in \mathbb{C} \,|\, |w| = 1\}$ è biunivoca, ossia $\forall \theta \in [0, 2\pi[$

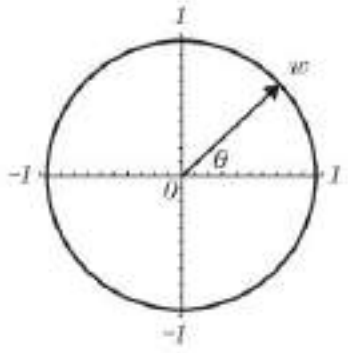


Figura 12.3: Rappresentazione trigonometrica di $w = \cos\theta + i\sin\theta$.

esiste ed è unico un $w \in \mathbb{C}$ con $|w| = 1$ tale che $w = \cos\theta + i\sin\theta$. Viceversa, fissiamo ora $\theta \in [0, 2\pi[$ e consideriamo $w = \cos\theta + i\sin\theta$, prendiamo un numero reale non negativo $\rho \in \mathbb{R}$, $\rho \geq 0$ e moltiplichiamolo scalarmente per w, otteniamo $z = \rho w = \rho(\cos\theta + i\sin\theta)$ che non è altro che un numero complesso che giace sulla retta individuata da O e da w ma che dista ρ da 0 anziché 1 come w (infatti si ha $|\rho w| = \sqrt{\rho^2\cos^2\theta + \rho^2\sin^2\theta} = |\rho|$). Facendo così variare θ in $[0, 2\pi[$ e ρ in tutto $\mathbb{R}^+$ otteniamo un'altra rappresentazione dei numeri complessi.

Definizione 12.4.1. I numeri complessi $\mathbb{C}$ possono essere descritti in *forma trigonometrica* nel modo seguente:

$$\mathbb{C} = \left\{ z = \rho(\cos\theta + i\sin\theta) \,|\, \rho \in \mathbb{R}^+, \theta \in [0, 2\pi[\right\}.$$

L'angolo θ viene chiamato *argomento* di z ed indicato con $arg(z)$; il numero reale positivo ρ è detto *modulo* $\rho = |z|$ di z.

Vediamo alcune banali relazioni tra la rappresentazione algebrica e quella trigonometrica.
Se $z \in \mathbb{C}$ si può scrivere come $z = a + ib = \rho(\cos\theta + i\sin\theta) \in \mathbb{C}$ allora

- $a = Re(z) = \rho\cos\theta$, $b = Im(z) = \rho\sin\theta$;
- $\rho = \sqrt{a^2 + b^2}$;
- Si ha:
$$\cos\theta + i\sin\theta = \frac{a + ib}{\sqrt{a^2 + b^2}},$$
quindi
$$\cos\theta = \frac{a}{\sqrt{a^2 + b^2}} \quad \text{e} \quad \sin\theta = \frac{b}{\sqrt{a^2 + b^2}}.$$

La forma trigonometrica è comoda per scrivere prodotti e quozienti tra i numeri complessi.

Siano $z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1)$ e $z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2)$, allora

$$z_1.z_2 = \rho_1(\cos\theta_1 + i\sin\theta_1).\rho_2(\cos\theta_2 + i\sin\theta_2) =$$

$$= \rho_1\rho_2(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2)) =$$

$$= \rho_1\rho_2(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \sin\theta_2)).$$

Da cui deduciamo il seguente teorema.

Teorema 12.4.1 (Teorema di De Moivre). *Siano* $z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1)$, $z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2) \in \mathbb{C}$, *allora*

$$|z_1.z_2| = \rho_1.\rho_2 = |z_1|.|z_2|;$$
$$arg(z_1.z_2) = arg(z_1) + arg(z_2).$$

Analogamente si vede che se $z_2 \neq 0$ allora

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}(\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)).$$

Da cui

$$\left|\frac{z_1}{z_2}\right| = \frac{\rho_1}{\rho_2} = \frac{|z_1|}{|z_2|}$$

$$arg\left(\frac{z_1}{z_2}\right) = arg(z_1) - arg(z_2).$$

A questo punto è abbastanza semplice dedurre che

$$z_1 \cdots z_n = \rho_1 \cdots \rho_n(\cos(\theta_1 + \cdots + \theta_n) + i\sin(\theta_1 + \cdots + \theta_n)).$$

Analogamente

$$z^n = \rho^n(\cos(n\theta) + i\sin(n\theta)).$$

Definizione 12.4.2. Se per due numeri complessi $z, w \in \mathbb{C}$ accade che

$$z^n = w$$

si dice che z è una radice n-esima di w.

Fissato un $w \in \mathbb{C}$ ci possiamo chiedere quante radici n-esime esso ammetta.

Teorema 12.4.2. *Sia $w \in \mathbb{C}$ ed $n \in \mathbb{Z}$ con $n \geq 1$, allora esistono esattamente n radici n-esime complesse $z_0, z_1, \ldots, z_{n-1}$ di w.*
Se $w = r(\cos\varphi + i\sin\varphi)$ e $z_k = \rho_k(\cos\theta_k + i\sin\theta_k)$, con $k = 0, \ldots, n-1$, si ha che

$$\rho_k = r^{1/n}, \quad \theta_k = \frac{\varphi + 2k\pi}{n}$$

per $k = 0, \ldots, n-1$.

Dimostrazione. Se $w = 0$ allora l'equazione $z^n = 0$ ha esattamente n soluzioni tutte coincidenti con zero. Quindi questo caso era banale e assumiamo ora che $w \neq 0$. Dobbiamo verificare che:

1. $z_k^n = w$, $\forall k = 0, \ldots, n-1$,
2. non vi siano altre soluzioni.

Iniziamo le verifiche.

1 $z_k^n(\rho_k(\cos\theta_k + i\sin\theta_k))^n$ per De Moivre risula essere uguale a
$\rho_k^n(\cos(n\theta_k) + i\sin(n\theta_k)) = (r^{1/n})^n \left(\cos\left(n \cdot \frac{\varphi+2k\pi}{n}\right) + i\sin\left(n \cdot \frac{\varphi+2k\pi}{n}\right)\right) =$
$= r(\cos(\varphi 2k\pi) + i\sin(\varphi + 2k\pi)) = r(\cos\varphi + i\sin\varphi) = w.$

1. Supponiamo esista $q = R(\cos\phi + i\sin\phi)$ tale che $q^n = w$. Ora
$q^n = R^n(\cos(n\phi) + i\sin(n\phi))$, quindi $R^n = r$ da cui $R = r^{1/n}$ perciò $R = \rho_k$.
Inoltre $n\phi = \varphi + 2h\pi$ con $h \in \mathbb{Z}$, dunque $\phi = \frac{\varphi+2h\pi}{n}$. Dando a h i valori da 0 a $n-1$ troviamo proprio gli z_k, dandone un qualunque altro, diciamo $\tilde{h}$, esso si potrà sempre scrivere come $\tilde{h} = k + mn$ per un certo $m \in \mathbb{Z}$ per cui sarebbe
$\phi = \frac{\varphi+2\tilde{h}\pi}{n} = \frac{\phi+2(k+mn)\pi}{n} = \frac{\phi+2k\pi}{n} + \frac{2mn\pi}{n} = \theta_k + 2m\pi$
e ritroviamo quindi gli stessi z_k precedenti.

□

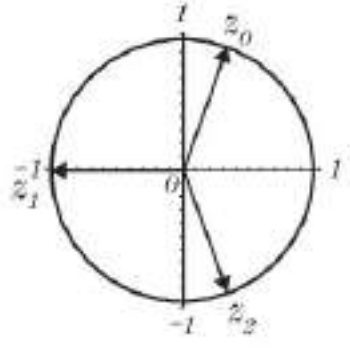


Figura 12.4: Rappresentazione trigonometrica delle radici cubiche del numero complesso $w = \cos\pi + i\sin\pi$.

Esempio 12.4.1. Cerchiamo le radici cubiche di $w = -1$. Scriviamo subito w in forma trigonometrica: $w = \cos\pi + i\sin\pi$. Quindi, le radici cercate sono

$z_0 = \cos(\pi/3) + i\sin(\pi/3) = \mathbf{1/2} + \sqrt{3}i/2,$
$z_1 = \cos(\pi) + i\sin(\pi) = \mathbf{-1},$
$z_2 = \cos(5\pi/3) + i\sin(5\pi/3) = 1/2 - \sqrt{3}i/2.$

Osserviamo la disposizione della radici cubiche nell'esempio precedente: esse si trovano ai vertici di un triangolo equilatero (vedi Figura 12.4). Questo fenomeno non è casuale infatti se $w = r(\cos\varphi + i\sin\varphi)$, le sue radici n-esime $z_0, \ldots, z_{n-1}$ si trovano ai vertici del poligono regolare di n lati inscritto nella circonferenza di centro O e raggio $r^{1/n}$ con vertice z_0 posto nel punto di angolo $\theta = \varphi/n$.

Esercizio 12.4.1. Calcolare e disegnare sul piano complesso le radici seste di i.

Osserviamo che se cercassimo le radici n-esime di 1, poiché si ha: $1 = \cos 0 + i\sin 0$, allora z_0 sarebbe sempre 1, infatti $arg(z_0) = 0/n = 0$. Da questa semplice osservazione si vede che le radici n-esime dell'unità individuano sempre un poliedro regolare di n lati con vertice in 1. Esse saranno pertanto del tipo:

$$z_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}$$

con $k = 0, \ldots, n-1$. Quindi

$$z_0 = \mathbf{1}, \quad \text{e} \quad z_1 = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n},$$

per $k = 2, \ldots, n-1$, si ha che $z_k = z_1^k$.

Perciò se chiamiamo $\omega := \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$, allora le radici n-esime dell'unità sono

$$1, \omega, \omega^2, \ldots, \omega^{n-1}.$$

Non è quindi difficile vedere che se $z^n = w = r(\cos\varphi + i\sin\varphi)$ qualunque sia $w \in \mathbb{C}$, le radici n-esime di w sono:

$$z, z\cdot\omega, z\cdot\omega^2, \ldots, z\cdot\omega^{n-1}$$

con $z = r^{1/n}(\cos\varphi/n + i\sin\varphi/n)$, infatti z è sicuramente una radice n-esima, per il teorema di De Moivre, se z_1 e z_2 sono due radici n-esime di w, allora $(z_1/z_2)^n = w/w = 1$, quindi z_1/z_2 è una radice n-esima di 1. Viceversa se z è una radice n-esima di w e ω è una radice n-esima di 1 allora $(z\cdot\omega)^n = w$.

Queste considerazioni sulle radici n-esime di un numero complesso, e in particolare quelle sulle radici dell'unità, permettono di affrontare con maggior consapevolezza la soluzione delle equazioni algebriche. L'equazione $x^2 + 1 = 0$ che abbiamo visto nell'introduzione di questo capitolo non ammette soluzioni nei reali, ma, riletta alla luce di quanto appena descritto nei complessi, corrisponde a trovare le radici di -1 che sono esattamente due ossia $\pm i$. Non solo, tutta questa descrizione ci dice anche che l'equazione $x^n = y$ ha sempre esattamente n soluzioni nei

complessi, qualunque sia $y \in \mathbb{C}$ (cosa falsa in $\mathbb{R}$). Quanto appena dimostrato è di portata assai più generale, vale infatti il seguente teorema che non dimostreremo.

Teorema 12.4.3 (Teorema fondamentale dell'algebra). *Un'equazione polinomiale della forma*

$$a_0 + a_1 x + \cdots + a_n x^n = 0$$

con $a_n \neq 0$ a coefficienti complessi qualsiasi ha precisamente n radici in $\mathbb{C}$ (se ognuna viene contata con la propria molteplicità).

12.5 Rappresentazione esponenziale

Diamo qui un'ultima possibile rappresentazione dei numeri complessi.

Consideriamo la circonferenza unitaria di centro l'origine: $S = \{z \in \mathbb{C} \,|\, |z| = 1\}$. Sia poi

$$\begin{array}{rcl} \Phi : \mathbb{R} & \to & S \\ t & \mapsto & (\cos t, \sin t) \end{array} .$$

Osservazioni:

- Φ non è iniettiva ed è periodica di periodo 2π;
- Se $z_1, z_2 \in S$, allora $z_1 . z_2 \in S$ (infatti $|z_1 . z_2| = |z_1| . |z_2| = 1$);
- Se $z \in S$ allora anche z^{-1} e $\overline{z}$ stanno su S;
- $1 \in S$;
- $\Phi(t_1 + t_2) = \Phi(t_1)\Phi(t_2)$.

Queste osservazioni ci fanno pensare che esista una parentela formale tra la funzione esponenziale e le funzioni trigonometriche, ma esse indicano anche che la parentela viene messa in luce nel campo complesso.

Definizione 12.5.1. *Se $z \in \mathbb{C}$ e $z = (x, y)$ definiamo la mappa esponenziale nel modo seguente:*

$$\begin{array}{rcl} \Phi_{\mathbb{C}} : \mathbb{C} & \to & \mathbb{C} \\ z & \mapsto & (e^x \cos y, e^x \sin y) \end{array} .$$

Osservazioni:

- $\Phi_C(0, y) = (\cos y, \sin y) = \Phi(y)$;
- $\Phi_C(x, 0) = (e^x, 0) =$ funzione esponenziale in base e;
- $\Phi_C(z_1 + z_2) = \Phi_C(z_1)\Phi_C(z_2)$ (lasciamo la verifica di questo semplice fatto per esercizio).

Per analogia con la nomenclatura nel caso reale, Φ_C viene chiamata *esponenziale complesso* e denotato con

$$\Phi_C(z) = \exp z \quad \text{o anche} \quad e^z,$$

si pone cioè , se $z = x + iy$:

$$e^z = e^x \cos y + ie^x \sin y.$$

Se $y = 0$ allora $e^z = e^x$; se invece $x = 0$, allora

$$e^{iy} = \cos y + i \sin y.$$

Quella appena scritta è la nota formula di Eulero e altro non è che la scrittura trigonometrica del numero complesso w di modulo 1 e argomento y.

Teorema 12.5.1 (Formula di Eulero). *Per ogni numero reale y si ha che*

$$e^{iy} = \cos y + i \sin y.$$

Quindi $e^z = e^x e^{iy} = e^x(\cos y + i \sin y)$ non è altro che la scrittura di un numero complesso di argomento y e modulo e^x. Possiamo perciò dare un'altra rappresentazione dei numeri complessi, ossia la *rappresentazione esponenziale* e scrivere:

Definizione 12.5.2. Un numero complesso z nella sua *rappresentazione esponenziale* si scrive

$$z = \rho e^{i\theta}$$

dove $\rho = |z|$ e $\theta = arg(z)$.

Esempio 12.5.1.

- $1 = e^0 = e^{2\pi i} = e^{2k\pi i}$, $\forall k \in \mathbb{Z}$;
- $-1 = e^{\pi i} = e^{-\pi i} = e^{(2k+1)\pi i}$ $\forall k \in \mathbb{Z}$; questo esempio prova anche una delle più famose formule in matematica in quanto contiene tutti i più importanti numeri della storia $(e, \pi, i, 1, 0)$:

$$e^{\pi i} + 1 = 0.$$

- $i = e^{\pi i/2} = e^{\pi i(1/2+2k)}$, $\forall\, k \in \mathbb{Z}$;
- $-i = e^{-\pi i/2 = i\pi(-1/2+2k)}$, $\forall\, k \in \mathbb{Z}$;
- In generale $e^{z+2k\pi i} = e^z$ $\forall\, k \in \mathbb{Z}$, quindi l'esponenziale complesso è una funzione $e^{i\theta}$ periodica di periodo 2π (rispetto alla variabile reale θ).

Osserviamo che $\forall\, z \in \mathbb{C}$, si ha $e^z \neq 0$; infatti, se supponiamo che sia $e^z = 0$, ossia che esista uno z tale che $e^z = e^x(\cos y + i \sin y) = 0$ allora

$$\begin{cases} e^x \cos y = 0 \\ e^x \sin y = 0 \end{cases} \text{ e quindi, dato che } e^x > 0, \quad \begin{cases} \cos y = 0 \\ \sin y = 0 \end{cases}$$

che è impossibile.

Ricaviamo ora un paio di formule importanti. Iniziamo con l'osservare che $e^{iy} = \cos y + i \sin y$ e $e^{-iy} = \cos(-y) + i \sin(-y) = \cos y - i \sin y$. Sommando queste due otteniamo che $e^{iy} + e^{-iy} = 2 \cos y$ da cui

$$\cos y = \frac{e^{iy} + e^{-iy}}{2}.$$

Se invece di sommare le due formule sopra le sottraiamo, otteniamo che $e^{iy} - e^{-iy} = 2i \sin y$ da cui

$$\sin y = \frac{e^{iy} - e^{-iy}}{2i}.$$

Possiamo usare queste due formule per dare delle definizioni.

Definizione 12.5.3. *Le funzioni trigonometriche a valori complessi vengono così definite per ogni* $z \in \mathbb{C}$*:*

$$\sin z := \frac{e^{iz} - e^{-iz}}{2i};$$

$$\cos z := \frac{e^{iz} + e^{-iz}}{2};$$

$$\operatorname{tg} z := \frac{e^{iz} - e^{-iz}}{i(e^{iz} + e^{-iz})}.$$

Si verifica facilmente (non lo facciamo qui) che tutte le proprietà formali note dalla trigonometria (periodicità, simmetria, formule di addizione, duplicazione, prostaferesi ecc.) restano valide.

12.6 Esercizi

Esercizio 12.6.1. Calcolare e disegnare sul piano complesso **tutti** gli $z \in \mathbb{C}$ tali che $|z| = 2$.

Esercizio 12.6.2. Calcolare e disegnare sul piano complesso **tutti** gli $z \in \mathbb{C}$ tali che $arg(z) = \pi/4$.

Esercizio 12.6.3. Disegnare sul piano complesso il punto $z \in \mathbb{C}$ tale che $|z| = 1$ e $arg(z) = 3\pi/2$.

Esercizio 12.6.4. Calcolare in modo esplicito le seguenti espressioni: $(3-i)^3$, $i^{19} - i^{11}$, $(-1)^{1/3}$, $(i)^{1/6}$.

Esercizio 12.6.5. Verificare la seguente uguaglianza: $(\overline{z}i + 1)/2 = z$.

Esercizio 12.6.6. Disegnare su piano complesso l'insieme:

$$\{z \in \mathbb{C} |\ |z + 2i| = |z - 1|\}.$$

Esercizio 12.6.7. Risolvere la seguente equazione $2z^2 + 5z - 1 = 0$.

Esercizio 12.6.8. Data l'equazione $z^2 + \lambda z - 2i(1+i) = 0$, determinare λ in modo tale che $z = 2 + i$ ne sia soluzione. Per tale valore di λ determinare l'altra radice.

Esercizio 12.6.9. Risolvere le disequazioni

$$z\overline{z} + 2z + 2\overline{z} < 4 \quad \text{e} \quad |z + 1| + i|z| = |i + z| + i.$$

Esercizio 12.6.10. Scrivere nella forma algebrica, trigonometrica ed esponenziale il numero $z \in \mathbb{C}$, corrispondente a $(1, \sqrt{3})$ in rappresentazione cartesiana.

Esercizio 12.6.11. Risolvere l'equazione $(27i - z^3)(z^2 - iz + 12) = 0$.

Esercizio 12.6.12. Risolvere l'equazione $i\cos\theta + \sin\theta = ie^{i\theta} + 1$.

Esercizio 12.6.13. Verificare che $\forall z \in \mathbb{C}$, si ha $\sin 2z = 2\sin z \cos z$.

CAPITOLO 13

Soluzione di alcuni esercizi

Proponiamo qui le soluzioni per alcuni esercizi scelti fra quelli proposti nei vari capitoli del libro. Ricordiamo che le soluzioni di tutti gli altri esercizi si possono trovare nel sito collegato al libro.

13.1 Soluzioni Capitolo 1

Soluzione Esercizio 1.6.6

La funzione $f(n) = 2n + 1$ è iniettiva in quanto $f(n_1) = f(n_2)$, implica $2n_1 + 1 = 2n_2 + 2$, da cui si ricava facilmente che $n_1 = n_2$. La funzione non è suriettiva come funzione tra numeri naturali $f : \mathbb{N} \to \mathbb{N}$, in quanto, per esempio, $0 \neq f(n), \forall n \in \mathbb{N}$.

Possiamo notare che se consideriamo $f : \mathbb{Q} \to \mathbb{Q}$, allora la f è suriettiva, in quanto per ogni $b \in \mathbb{Q}$, si ha $f(x) = b$ per $x = \frac{b-1}{2}$.

Soluzione Esercizio 1.6.11

Non è difficile notare che l'operazione $*$ è commutativa, visto che la somma e il prodotto lo sono.

Per vedere che l'operazione $*$ è associativa, consideriamo, $\forall a, b, c \in \mathbb{Q}$:

$$a * (b * c) = a * (bc + b + c) = a(bc + b + c) + a + bc + b + c =$$

$$= abc + ab + ac + bc + a + b + c;$$

ma anche

$$(a * b) * c = (a * b)c + (a * b) + c = (ab + a + b)c + ab + a + b + c =$$

$$= abc + ac + bc + ab + a + b + c,$$

quindi la proprietà associativa vale.

Si ha poi che 0 è l'elemento neutro, in quanto, $\forall\, x \in \mathbb{Q}$:

$$0 * a = 0 \cdot a + a + 0 = a.$$

Dobbiamo infine vedere se ogni elemento $x \in \mathbb{Q}$ ammette l'inverso rispetto a $*$, dobbiamo cioè vedere se

$\forall\, x \in \mathbb{Q}$, $\exists\, y \in \mathbb{Q}$, tale che $x * y = 0$.

Dovremmo quindi avere:

$$x * y = xy + x + y = 0, \quad \text{quindi} \quad y(x+1) = -x \quad \text{e} \quad y = \frac{-x}{x+1};$$

ma allora ogni elemento di x ha inverso, pari a $\frac{-x}{x+1}$, tranne però $x = -1$, per cui questo valore non ha senso. Quindi $(\mathbb{Q}, *)$ non è un gruppo commutativo (anche se ci era andato vicino!).

In effetti il solo problema è che l'elemento -1 non ha inverso. Se consideriamo l'insieme $\mathbb{Q}' = \mathbb{Q} - \{-1\}$, la struttura $(\mathbb{Q}', *)$ è un gruppo commutativo.

13.2 Soluzioni Capitolo 2

Soluzione Esercizio 2.5.2

La prima proprietà che un sottoinsieme deve verificare per essere un sottospazio vettoriale è banalmente vera per $\mathcal{V}_L$ in quanto per ipotesi L passa per l'origine, quindi il vettore nullo sta in $\mathcal{V}_L$.

La seconda proprietà richiede che per ogni $\alpha, \beta \in \mathbb{R}$ e per ogni $\overrightarrow{\mathbf{w}}_1, \overrightarrow{\mathbf{w}}_2 \in \mathcal{V}_L$, si abbia $\alpha\overrightarrow{\mathbf{w}}_1 + \beta\overrightarrow{\mathbf{w}}_2 \in \mathcal{V}_L$.

Nel nostro caso se $\overrightarrow{\mathbf{w}}_1 \subset L$ allora $\alpha\overrightarrow{\mathbf{w}}_1$ starà ancora su L in quanto la moltiplicazione di un vettore per uno scalare non cambia la direzione del vettore; quindi $\alpha\overrightarrow{\mathbf{w}}_1 \in \mathcal{V}_L$ (analogamente $\beta\overrightarrow{\mathbf{w}}_2 \in \mathcal{V}_L$). Ora abbiamo che $\alpha\overrightarrow{\mathbf{w}}_1, \beta\overrightarrow{\mathbf{w}}_2 \subset L$ ossia $\alpha\overrightarrow{\mathbf{w}}_1$ e $\beta\overrightarrow{\mathbf{w}}_2$ sono allineati. La somma di due vettori allineati giace ancora sulla medesima retta, quindi $\alpha\overrightarrow{\mathbf{w}}_1 + \beta\overrightarrow{\mathbf{w}}_2 \in \mathcal{V}_L$.

Soluzione Esercizio 2.5.19

Il sottoinsieme W_1 è un sottospazio di $\mathbb{R}^4$? Dobbiamo verificare che valgano 1) e 2) della Proposizione 2.4.1, ove:

1) Se $a = b = 0$ allora la quaterna $(a, b, 2a, -b) = (\mathbf{0}, 0, \mathbf{0}, 0)$, quindi $0_{\mathbb{R}^4} \in W_1$.

2) Siano poi $\mathbf{v}_1, \mathbf{v}_2 \in W_1$, dunque esistono $a_1, a_2, b_1, b_2 \in \mathbb{R}$ tali che $\mathbf{v}_1 = (a_1, b_1, 2a_1, -b_1)$ e $\mathbf{v}_2 = (a_2, b_2, 2a_2, -b_2)$.

   Per ogni $\alpha, \beta \in \mathbb{R}$ si avrà :

$$\alpha\mathbf{v}_1 + \beta\mathbf{v}_2 = \alpha(a_1, b_1, 2a_1, -b_1) + \beta(a_2, b_2, 2a_2, -b_2) =$$

$$= (\alpha a_1 + \beta a_2, \alpha b_1 + \beta b_2, 2(\alpha a_1 + \beta a_2), -(\alpha b_1 + \beta b_2)) \in W_1.$$

Quindi W_1 è un sottospazio di $\mathbb{R}^4$.

Abbiamo poi che una base di W_1 è data da $\{(1,0,2,0),(0,1,0,1)\}$ in quanto ogni vettore appartenente a W_1 è del tipo $(a,b,2a,-b)$ e quindi può sempre essere scritto con la seguente combinazione lineare dei vettori suddetti: $a(1,0,2,0)+b(0,1,0,-1)$.

Questo prova che $\{(1,0,2,0),(0,1,0,1)\}$ è un sistema di generatori per W_1.

Inoltre $(a,b,2a,-b)=(0,0,0,0)$ se e solo se $a=b=0$, quindi $(1,0,2,0)$ e $(0,1,0,1)$ sono indipendenti, e formano una base.

Il sottoinsieme W_2 non è un sottospazio di $\mathbb{R}^4$ in quanto $\mathbf{0}\notin W_2$.

Il sottoinsieme W_3 è un sottospazio di $\mathbb{R}^4$ e la dimostrazione è analoga a quella svolta per W_1.

Il sottoinsieme W_4 non è un sottospazio di $\mathbb{R}^4$. Ad esempio i vettori $(2,0,2,4),(5,0,3,15)\in W_4$, ma

$$3(2,0,2,4)+2(5,0,3,15)=(16,0,12,42)\notin W_4.$$

Soluzione Esercizio 2.5.20

$V=\{(x,y,z)\in\mathbb{R}^3 \mid x-ky+z=0,\ kx-y+z=0\}$ è un sottospazio di $\mathbb{R}^3$ e dobbiamo calcolarne la dimensione al variare di $k\in\mathbb{R}$. Dobbiamo quindi cercare "il numero di soluzioni", al variare di $k\in\mathbb{R}$, del sistema lineare:

$$\begin{cases} x-ky+z=0 \\ kx-y+z=0 \end{cases}.$$

Si può verificare semplicemente (per esempio risolvendo il sistema per sostituzione, o tramite un'eliminazione di Gauss) che vi sono due casi da distinguere:

$k=1$. In questo caso le due equazioni sono la medesima: $x-y+z=0$, perciò le soluzioni del sistema sono date da $x=y-z$, e quindi $V=\{(x,y,z)\in\mathbb{R}^3 \mid x=y-z\}$ e gli elementi di V sono le terne del tipo $(y-z,y,z)$. Una base di V è data da $\{(1,1,0),(-1,0,1)\}$ (ogni elemento di V si scrive come $y(1,1,0)+z(-1,0,1)$ e i due vettori sono indipendenti) e dunque in questo caso $\dim V=2$.

$k\neq 1$. In questo caso il sistema diventa

$$\begin{cases} x=-y \\ z=-x(k+1) \end{cases}.$$

Perciò $V=\{(x,y,z)\in\mathbb{R}^3 \mid y=-x, z=-x(k+1)\}$, e gli elementi di V sono tutti del tipo $(x,-x,-x(k+1))$. Una base di V è data da $\{(1,-1,-k-1)\}$. Da ciò deduciamo semplicemente che in questo caso $\dim V=1$.

Soluzione Esercizio 2.5.22

Osserviamo innanzitutto che V è lo spazio dei polinomi in una variabile senza termine noto, quindi banalmente $0_{\mathbb{R}[x]} \in V$. Altrettanto banalmente la combinazione lineare di due polinomi in V è ancora un polinomio in V, infatti se $p_1(0) = p_2(0) = 0$ allora $(\alpha p_1)(0) = \alpha p_1(0) = \alpha \cdot 0 = 0$ (idem per $\beta p_2(0) = 0$), dunque

$$(\alpha p_1 + \beta p_2)(0) = (\alpha p_1)(0) + (\beta p_2)(0) = \alpha p_1(0) + \beta p_2(0) =$$

$$= \alpha \cdot 0 + \beta \cdot 0 = 0 + 0 = 0.$$

Mostriamo ora che $\dim V = \infty$. Se così non fosse esisterebbe un certo numero finito, diciamo $n \in \mathbb{N}$, di polinomi $p_1, \ldots, p_n \in V$ tali che $\langle p_1, \ldots, p_n \rangle = V$. Questi polinomi $p_1, \ldots, p_n$ avranno ciascuno un loro grado $d_1, \ldots, d_n$. Supponiamo che $d_n \geq d_i$ per ogni $i = 1, \ldots, n-1$ (ovviamente non cambia nulla se supponiamo che sia un altro a essere il maggiore, l'importante è che ne esisterà almeno uno). Dunque ogni $p \in V$ si scriverebbe come combinazione lineare di $p_1, \ldots p_n$, e p non potrebbe avere grado maggiore di d_n, e questo è un assurdo (per esempio i polinomi $x^k \in V$, $\forall\, k \geq 0$). L'assurdo deriva appunto dall'aver supposto che la base di V fosse finita.

Soluzione Esercizio 2.5.25

Il sottospazio $S \cap S'$ è dato da

$$S \cap S' = \{(x, y, z) \in \mathbb{R}^3 \mid x - y = 0, x + 2y - z = 0\}.$$

Dunque le coordinate di un elemento di tale sottospazio dovranno soddisfare il seguente sistema lineare:

$$\begin{cases} x - y = 0 \\ x + 2y - z = 0 \end{cases}$$

che ha soluzione

$$\begin{cases} x = y \\ z = 3y \end{cases}$$

perciò gli elementi di $S \cap S'$ sono tutti del tipo $(y, y, 3y)$ al variare di $y \in \mathbb{R}$. Una base di $S \cap S'$ è quindi data da $\{(1, 1, 3)\}$ (e $\dim(S \cap S') = 1$).

13.3 Soluzioni Capitolo 3

Soluzione Esercizio 3.5.5

Il terzo prodotto dell'Esercizio 3.5.1 ci mostra come dovrebbe essere fatta l'inversa di una matrice simmetrica

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \in \mathbb{R}^{2,2}.$$

Ma tale matrice esisterà solo se $a^2 - b^2 \neq 0$, quindi non per tutte le matrici simmetriche. In effetti la matrice A avrà inversa se e solo se $a^2 - b^2 \neq 0$, che è il *determinante* di A (vedi il Capitolo 6).

Soluzione Esercizio 3.5.8

$$A^2 - A - 8I = \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 3 & -1 \end{pmatrix} - 8\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 10 & 2 \\ 3 & 7 \end{pmatrix} + \begin{pmatrix} -2 & -2 \\ -3 & +1 \end{pmatrix} + \begin{pmatrix} -8 & 0 \\ 0 & -8 \end{pmatrix} =$$

$$= \begin{pmatrix} 10 & 2 \\ 3 & 7 \end{pmatrix} + \begin{pmatrix} -10 & -2 \\ -3 & -7 \end{pmatrix} = O.$$

Soluzione Esercizio 3.5.7:

Per ipotesi $A^2 = I$. Moltiplichiamo entrambi i membri dell'uguaglianza per A^{-1}. Otteniamo $A^{-1}A^2 = A^{-1}I$, che dà : $A^{-1} = A$.

Osservazione 1: Si è scelto di moltiplicare per A^{-1} a sinistra. Si poteva scegliere di farlo anche a destra?

Osservazione 2: Abbiamo usato il fatto A^{-1} esista. Se nel testo dell'esercizio non fosse scritto che A è invertibile, potremmo dedurre che se $A^2 = I$ allora necessariamente A è invertibile? In effetti è possibile, come potremo vedere usando il concetto di determinante (vedi Capitolo 6).

Soluzione Esercizio 3.5.9

La scrittura $AB = O$, $BA = O$ è un po' ambigua; infatti indichiamo con O la matrice nulla di ogni forma, ma in questo caso abbiamo $AB \in \mathbb{R}^{m,m}$ e $BA \in \mathbb{R}^{n,n}$; potremmo più correttamente scrivere: $AB = O_{m,m}$, $BA = O_{n,n}$ e quindi, se $m \neq n$, avremo $AB \neq BA$.

Soluzione Esercizio 3.5.32

Basta osservare che:

$$(AB)A = A \cdot A = A^2 \quad \text{e} \quad A(BA) = AB = A,$$

Poiché il prodotto fra matrici è associativo, si ha $(AB)A = A(BA)$, e quindi $A = A^2$ (in modo analogo si ha anche $B = B^2$).

13.4 Soluzioni Capitolo 4

Soluzione Esercizio 4.6.6

Dobbiamo considerare la matrice $A|B$, ove:

$$\left(\begin{array}{ccc|c} 1 & 2 & k & k+1 \\ 1 & k+3 & & 3 \\ 1 & 2 & k & k+1 \end{array}\right)$$

Eseguendo una eliminazione di Gauss, avremo:

$$\left(\begin{array}{ccc|c} 1 & 2 & k & k+1 \\ 1 & k+3 & & 3 \\ 1 & 2 & k & k+1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & k & k+1 \\ 0 & k+1 & -k & 2-k \\ 0 & 0 & 0 & 0 \end{array}\right);$$

si ha quindi che $\forall\, k \in \mathbb{R}$, $r(A) = r(A|B) = 2$, in quanto nella seconda riga ho il pivot $k+1$ se $k \neq -1$, e per $k = -1$ ho il pivot $-k = 1$. Quindi il sistema ha sempre $\infty^{3-2} = \infty^1$ soluzioni.

In particolare, avremo $(k+1)y = kz + 2 - k$, allora, per $k \neq -1$:

$$y = \frac{k}{k+1}z + \frac{2-k}{k+1}, \quad \text{e quindi} \quad x = -\frac{k(k+3)}{k+1}z + \frac{k^2+4k-3}{k+1};$$

e le soluzioni sono le terne: $(-\frac{k(k+3)}{k+1}z + \frac{k^2+4k-3}{k+1}, \frac{k}{k+1}z + \frac{2-k}{k+1}, z)$.

Invece per $k = -1$ si ha:

$$z = 3, \quad \text{da cui :} \quad x = -2y + 3;$$

e le soluzioni sono le terne: $(2y+3, y, 3)$.

Soluzione Esercizio 4.6.11

Dobbiamo considerare la matrice $A|B$, ove (abbiamo invertito di posto le prime due equazioni, per comodità):

$$A|B = \left(\begin{array}{ccc|c} 1 & 1 & h & 0 \\ 1 & h & 1 & 2h-1 \\ h & 1 & 1 & 5 \end{array}\right).$$

Eseguendo una eliminazione di Gauss, avremo:

$$\left(\begin{array}{ccc|c} 1 & 1 & h & 0 \\ 0 & h-1 & 1-h & 2h-1 \\ 0 & 1-h & 1-h^2 & 5 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|c} 1 & 1 & h & 0 \\ 0 & h-1 & 1-h & 2h-1 \\ 0 & 0 & 2-h-h^2 & 2h+4 \end{array}\right);$$

dobbiamo quindi considerare quando $h-1 = 0$ e/o $2-h-h^2 = 0$.

Si ha che $h-1=0$ per $h=1$, e la matrice diviene:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 6 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

e quindi $r(A)=1$ e $r(A|B)=2$; allora, poiché $r(A)\neq r(A|B)$, il sistema non ha soluzioni.

Studiando l'equazione $2-h-h^2=0$, si ha che le sue soluzioni sono $h=1$ (già considerata) e $h=-2$. Per $h=-2$ la matrice diventa:

$$\left(\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 0 & -3 & 3 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right);$$

quindi $r(A)=r(A|B)=2$ e il sistema ha $\infty^{3-2}=\infty^1$ soluzioni.

Infine, per $h\neq 1$ e $h\neq -2$, avremo che $r(A)=r(A|B)=3$, e quindi il sistema ha un'unica soluzione.

Soluzione Esercizio 4.6.17

Dobbiamo considerare la matrice $A|B$, ove:

$$A|B=\left(\begin{array}{cccc|c} 1 & 1 & -k & 4 & 5 \\ k & k & -3 & 1 & 3 \\ 3 & 3 & 5 & 3 & -k \end{array}\right).$$

Eseguendo una eliminazione di Gauss, avremo:

$$\left(\begin{array}{cccc|c} 1 & 1 & -k & 4 & 5 \\ 0 & 0 & k^2-3 & 1-4k & 3-5k-k-15 \\ 0 & 0 & 5+3k & -9 & \end{array}\right) \to$$

ponendo $k^2-3\neq 0$, cioè $k\neq \pm\sqrt{3}$, si ha:

$$\to \left(\begin{array}{cccc|c} 1 & 1 & -k & 4 & 5 \\ 0 & 0 & k^2-3 & 1-4k & 3-5k \\ 0 & 0 & 0 & -\frac{5+3k}{k^2-3}(1-4k)-9 & -\frac{(5+3k)(3-5k)}{k^2-3}-k-15 \end{array}\right);$$

$$\to \left(\begin{array}{cccc|c} 1 & 1 & -k & 4 & 5 \\ 0 & 0 & k^2-3 & 1-4k & 3-5k \\ 0 & 0 & 0 & -\frac{3k^2+17k+22}{k^2-3} & -\frac{(5+3k)(3-5k)}{k^2-3}-k-15 \end{array}\right)$$

Nell'ultima riga, avremo che $3k^2+17k+22=0$ (risolvendo l'equazione) per $k=-2$ o per $k=-\frac{11}{3}$. Analizziamo cosa accade in questi due casi:

Per $k=-2$, si ha che tutta l'ultima riga della matrice si annulla, quindi $r(A)=r(A|B)=2$ e il sistema ha $\infty^{4-2}=\infty^2$ soluzioni.

Per $=-\frac{11}{3}$, si ha che l'ultima riga di A si annulla, ma il termine $-\frac{(5+3k)(3-5k)}{k^2-3}-k-15\neq 0$, quindi $r(A)=2\neq r(A|B)=3$, e quindi il sistema non ha soluzioni.

Quando invece si ha $k \neq -2$ e per $k \neq -\frac{11}{3}$ (e anche $k \neq \pm\sqrt{3}$), avremo che $r(A) = r(A|B) = 3$, quindi il sistema ha $\infty^{4-3} = \infty^1$ soluzioni.

Ci resta da vedere cosa accade per $k = \pm\sqrt{3}$; in questo caso dobbiamo considerare la matrice all'inizio della eliminazione e sostituire $k \neq \pm\sqrt{3}$; otterremo:

$$\left(\begin{array}{cccc|c} 1 & 1 & -k & 4 & 5 \\ 0 & 0 & 0 & 1 \pm 4\sqrt{3} & 3 \pm 5\sqrt{3} \\ 0 & 0 & 5 \pm 3\sqrt{3} & -9 & \pm\sqrt{3} - 15 \end{array}\right) \rightarrow$$

$$\rightarrow \left(\begin{array}{cccc|c} 1 & 1 & -k & 4 & 5 \\ 0 & 0 & 5 \pm 3\sqrt{3} & -9 & \pm\sqrt{3} - 15 \\ 0 & 0 & 0 & 1 \pm 4\sqrt{3} & 3 \pm 5\sqrt{3} \end{array}\right)$$

ove è facile notare che $r(A) = r(A|B) = 3$, quindi il sistema, come nel caso generale, ha $\infty^{4-3} = \infty^1$ soluzioni.

Soluzione Esercizio 4.6.23

Dobbiamo considerare

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid x + 3y - 3z = -7,\ 3x + y = 0\}.$$

Gli elementi di S sono le soluzioni del sistema:

$$\begin{cases} x + 3y - 3z = -7 \\ 3x + y = 0 \end{cases}$$

Possiamo risolvere il sistema per sostituzione:

$$\begin{cases} y = -3x \\ x + 3(-3x) - 3z = -7 \end{cases} \rightarrow \begin{cases} y = -3x \\ -3z = 8x - 7 \end{cases} \rightarrow \begin{cases} y = -3x \\ z = \frac{7-8x}{3} \end{cases}$$

Quindi gli elementi di S si possono scrivere come: $(x, -3x, \frac{7-8x}{3})$.

Si vede facilmente che $S = \{(x, -3x, \frac{7-8x}{3}) |\ x \in \mathbb{R}\}$ non può essere un sottospazio di $\mathbb{R}^3$, in quanto $(0, 0, 0) \notin S$.

Soluzione Esercizio 4.6.27

Considerando i due sottospazi di $\mathbb{R}^4$ definiti da:

$$S_1 := \{(x, y, z, t) \in V \mid hy + t = 0,\ 2x - z = 0\},$$

$$S_2 := \{(x, y, z, t) \in V \mid y - hz = 0,\ (h+1)x - 2z - ht = 0\}.$$

Avremo che gli elementi di $S_1 \cap S_2 \subseteq \mathbb{R}^4$ sono le quaterne (x, y, z, t) che soddisfano sia le equazioni che definiscono S_1 che quelle di S_2; quindi $S_1 \cap S_2$ è definito dal sistema:

$$\begin{cases} hy + t = 0 \\ 2x - z = 0 \\ y - hz = 0 \\ (h+1)x - 2z - ht = 0 \end{cases}$$

Dobbiamo considerare la matrice A, ed eseguire l'eliminazione di Gauss su di essa (abbiamo cambiato di posto le equazioni, per comodità):

$$A = \begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & -h & 0 \\ 0 & h & 0 & 1 \\ h+1 & 0 & -2 & -h \end{pmatrix} \to \begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & -h & 0 \\ 0 & h & 0 & 1 \\ 0 & 0 & \frac{h+1}{2} - 2 & -h \end{pmatrix} \to$$

$$\to \begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & -h & 0 \\ 0 & 0 & h^2 & 1 \\ 0 & 0 & \frac{h-3}{2} & -h \end{pmatrix} \to \text{(ponendo } h \neq 0\text{:)}$$

$$\to \begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & -h & 0 \\ 0 & 0 & h^2 & 1 \\ 0 & 0 & 0 & -\frac{h-3}{2h^2} - h \end{pmatrix} \to \begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & -h & 0 \\ 0 & 0 & h^2 & 1 \\ 0 & 0 & 0 & -\frac{-2h^3-h+3}{2h^2} \end{pmatrix}.$$

È facile vedere che $-2h^3 - h + 3$ si annulla per $h = 1$; si avrà (dividendo) $-2h^3 - h + 3 = (h-1)(2h^2 + 2h + 3)$, e $2h^2 + 2h + 3$ non ha radici reali; quindi l'unico valore critico ottenuto è $h = 1$, e si ha:

Per $h = 1$, A diviene:

$$\begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

quindi $r(A) = 3$, e il sistema ha $\infty^{4-3} = \infty^1$ soluzioni, cioè avremo che $\dim(S_1 \cap S_2) = 1$.

Se invece $h \neq 1$ (e $h \neq 0$), abbiamo $r(A) = 4$, quindi il sistema ha solo la soluzione nulla, cioè $\dim(S_1 \cap S_2) = 0$.

Rimane da vedere cosa accade per $h = 0$. In questo caso dobbiamo sostituire $h = 0$ nella matrice, ottenendo:

$$\begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & -h & 0 \\ 0 & 0 & h^2 & 1 \\ 0 & 0 & \frac{h-3}{2} & -h \end{pmatrix} \to \begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & -0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{-3}{2} & 0 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & -0 & 0 \\ 0 & 0 & \frac{-3}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Poiché di nuovo si ha $r(A) = 4$, anche qui $\dim(S_1 \cap S_2) = 0$, come nel caso generale.

Soluzione Esercizio 4.6.30

Un sistema di 3 equazioni in x, y avente $(3, 0)$ come unica soluzione può essere, ad esempio:

$$\begin{cases} x = 3 \\ y = 0 \\ 2x + 3y = 6 \end{cases}$$

Perché il sistema abbia un'unica soluzione si ha: $r(A) = r(A|B) = 2$, quindi la matrice $A|B$ (che è una 3×3) dovrà avere che una sua riga è combinazione lineare delle altre 2 (avendo rango 2).

Non è possibile che le tre equazioni (e cioè le tre righe di $A|B$) siano proporzionali (per esempio nel caso dato sopra ciò non accade); se accadesse avremmo $r(A|B) = 1$ e quindi il sistema avrebbe $\infty^{2-1} = \infty^1$ soluzioni, e non una sola.

Soluzione Esercizio 4.6.31

Dobbiamo costruire, se possibile, dei sistemi di 2 equazioni in 3 incognite, tali che:

1. il sistema ammetta una e una sola soluzione;

   È impossibile: se il sistema è $AX = B$, avremo che $r(A) \leq 2$, cosicché il sistema avrà $\infty^{3-r(A)}$ soluzioni, e quindi almeno ∞^1 soluzioni.

2. esistano 10 diverse soluzioni del sistema;

   Basta costruire un sistema che ammetta infinite soluzioni (ne esisteranno quindi ben più di 10). Un esempio (nelle incognite x, y, z):

   $$\begin{cases} x + y = 0 \\ y - z = 0 \end{cases}$$

3. il sistema ammetta esattamente 4 soluzioni;

   È impossibile; un sistema lineare o ha infinite soluzioni o una sola o nessuna.

4. il sistema non ammetta soluzioni.

   Ad esempio (nelle incognite x, y, z):

   $$\begin{cases} x + y = 0 \\ x + y = 6 \end{cases}$$

   Le due equazioni sono ovviamente incompatibili.

13.5 Soluzioni Capitolo 5

Soluzione Esercizio 5.6.1

1. La f non è lineare; per esempio $f(1) = 3$, mentre $f(2) = 8 \neq 2f(1)$.

2. $\forall x \in R$, si ha $f(x) = 1 - 1 = 0$; quindi la f è la funzione nulla, che abbiamo visto essere lineare.

3. La f è lineare: infatti $\forall (a,b),(c,d) \in \mathbb{R}^2$ avremo:

$$f((a,b)+(c,d)) = f(a+c,b+d) = (a+c+b+d, a+c-b-d) =$$

$$= (a+b, a-b) + (c+d, c-d) = f(a,b) + f(c,d);$$

e anche:

$$f(\lambda a, \lambda b) = (\lambda a + \lambda b, \lambda a - \lambda b) = \lambda(a+b, a-b) = \lambda f(a,b).$$

4. La f non è lineare; per esempio $f(1,1,1) = (1,1,3)$, ma $f(2,2,2) = (2,2,12) \neq 2f(1,1,1)$.

5. La f non è lineare; per esempio $f(1,1,1) = (1,1,1)$, ma $f(2,2,2) = (2,4,8) \neq 2f(1,1,1)$.

6. La f è lineare; infatti $\forall (a,b,c),(x,y,z) \in \mathbb{R}^3$ avremo:

$$f((a,b,c)+(x,y,z)) = f(a+x,b+y,c+z) = a+x+b+y+c+z =$$

$$= (a+b+c) + (x+y+z) = f(a,b,c) + f(x,y,z);$$

e anche:

$$f(\lambda a, \lambda b, \lambda c) = \lambda a + \lambda b + \lambda c = \lambda(a+b+c) = \lambda f(a,b,c).$$

7. La f non è lineare; per esempio $f(0,0,0) = 1$, ma dovrebbe essere $= 0$ se la f fosse lineare.

Soluzione Esercizio 5.6.3

Basta osservare che $(0,1,1) + (2,0,1) = (2,1,2)$, ma $f(0,1,1) + f(2,0,1) = (3,0,0) + (2,1,2) = (5,1,2) \neq f(2,1,2)$.

Soluzione Esercizio 5.6.4

Per scrivere la matrice $M_{\mathcal{E}_3,\mathcal{E}_3}(f)$ ci manca solo l'espressione esplicita di $f(\mathbf{e}_3)$, che è :

$$f(\mathbf{e}_3) = f(\mathbf{e}_1) - 2f(\mathbf{e}_2) = (2,1,-3) - 2(1,3,-2) = (0,-5,1).$$

Quindi avremo:

$$A = M_{\mathcal{E}_3,\mathcal{E}_3}(f) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -5 \\ -3 & -2 & 1 \end{pmatrix}.$$

Per calcolare $r(A)$, eseguiamo l'eliminazione di Gauss, dopo aver invertito le prime due righe, per comodità:

$$\begin{pmatrix} 1 & 3 & -5 \\ 2 & 1 & 0 \\ -3 & -2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -5 \\ 0 & -5 & 10 \\ 0 & 7 & -14 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -5 \\ 0 & -5 & 10 \\ 0 & 0 & 0 \end{pmatrix}.$$

Si ha allora che $r(A) = 2 = \dim \operatorname{Im} f$, perciò la f non è suriettiva. Inoltre $\dim \ker f = 3 - 2 = 1$ e la f non è nemmeno iniettiva.

Per verificare che $\mathcal{B}$ è una base di $\mathbb{R}^3$ mettiamo i suoi vettori come righe di una matrice e controlliamo (con un'eliminazione di Gauss) che il suo rango sia 3, cosicché essi siano linearmente indipendenti. In effetti si ha:

$$\begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -2 \end{pmatrix}$$

e quindi $\mathcal{B}$ è una base.

Per determinare la matrice $M_{\mathcal{E}_3,\mathcal{B}}(f)$ dobbiamo esprimere i vettori $f(\mathbf{e}_i)$, $i = 1, 2, 3$, tramite la base $\mathcal{B}$, quindi:

$$f(\mathbf{e}_1) = (2, 1, -3) = a\mathbf{v}_1 + n\mathbf{v}_2 + c\mathbf{v}_3 = (3b, a + 2b + c, 2a + b)$$

che ci dà il sistema:

$$\begin{cases} 3b = 2 \\ a + 2b + c = 1 \\ 2a + b = -3 \end{cases} \to \begin{cases} b = \frac{2}{3} \\ 2a = \frac{-9-2}{3} = -\frac{11}{3} \\ -\frac{11}{6} + \frac{4}{3} + c = 1 \end{cases} \to \begin{cases} b = \frac{2}{3} \\ a = -\frac{11}{6} \\ c = 1 + \frac{11}{6} - \frac{4}{3} = \frac{3}{2} \end{cases}$$

Quindi la prima colonna di $M_{\mathcal{E}_3,\mathcal{B}}(f)$ è $\left(-\frac{11}{6}, \frac{2}{3}, \frac{3}{2}\right)$; analogamente avremo per la seconda colonna:

$$\begin{cases} 3b = 1 \\ a + 2b + c = 3 \\ 2a + b = -2 \end{cases}.$$

Che dà : $(a, b, c) = \left(-\frac{7}{6}, \frac{1}{3}, \frac{7}{2}\right)$. Infine per la terza colonna avremo:

$$\begin{cases} 3b = 0 \\ a + 2b + c = -5 \\ 2a + b = 1 \end{cases}$$

Che dà : $(a, b, c) = \left(\frac{1}{2}, 0, -\frac{11}{2}\right)$. Si ha quindi:

$$M_{\mathcal{E}_3,\mathcal{B}}(f) = \begin{pmatrix} -\frac{11}{6} & -\frac{7}{6} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{3}{2} & \frac{7}{2} & -\frac{11}{2} \end{pmatrix}.$$

Soluzione Esercizio 5.6.5

Consideriamo a matrice $A = M_{\mathcal{E}_3,\mathcal{E}_3}(f)$ e eseguiamo l'eliminazione di Gauss:

$$A = \begin{pmatrix} m & 4 & -1 \\ 1 & 2 & -1 \\ 0 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 \\ 0 & 4 - 2m & m - 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Da qui si ha che se $4 - 2m \neq 0$, cioè $m \neq 2$, allora $r(A) = 3 = \dim \operatorname{Im} f$ e quindi la f è suriettiva. Poiché dominio e codominio hanno la stessa

dimensione, la f è anche iniettiva ($\dim \ker f = n - r(A) = 3 - 3 = 0$); così la f è un isomorfismo.

Se invece $m = 2$, allora:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix} \to A' = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

e $r(A) = 2 = \dim \mathrm{Im} f$; quindi la f non è suriettiva e neanche iniettiva (infatti $\dim \ker f = n - r(A) = 3 - 2 = 1$).

Dobbiamo ora determinare, per $m = 2$, una base per $\mathrm{Im} f$ e una per $\ker f$. Nel caso di $\mathrm{Im} f$, una base sarà data da due colonne indipendenti di A, per esempio: $\{(1,0,0),(-1,1,2)\}$. Per trovare una base di $\ker f$ dovremo risolvere il sistema $A \cdot X = O$, equivalente al sistema $A' \cdot X = O$, ove $X = {}^t(x,y,z)$:

$$\begin{cases} x + 2y - z = 0 \\ z = 0 \\ 0 = 0 \end{cases} \to \begin{cases} x = -2y \\ z = 0 \end{cases}.$$

Quindi $\ker f = \{(-2y, y, 0)\}$, e una sua base è : $\{(-2,1,0)\}$.

Soluzione Esercizio 5.6.6

Le due colonne della matrice cercata sono date da $id_{\mathbb{R}^2}(\mathbf{v}_1) = \mathbf{v}_1$ e $id_{\mathbb{R}^2}(\mathbf{v}_2) = \mathbf{v}_2$, ma scritti rispetto alla base di arrivo $\mathcal{B}'$; quindi dovremo determinare a, b in modo che:

$(2,1) = a(1,-1) + b(1,1)$ e $(0,1) = a(1,-1) + b(1,1)$;

nel primo caso avremo:

$$\begin{cases} 2 = a + b \\ 1 = -a + b \end{cases} \to \begin{cases} a = 2 - b \\ 2b - 2 = 1 \end{cases} \to \begin{cases} a = \frac{1}{2} \\ b = \frac{3}{2} \end{cases}.$$

Mentre nel secondo:

$$\begin{cases} 0 = a + b \\ 1 = -a + b \end{cases} \to \begin{cases} a = -b \\ 2b = 1 \end{cases} \to \begin{cases} a = -\frac{1}{2} \\ b = \frac{1}{2} \end{cases}.$$

Quindi la matrice cercata è:

$$M_{\mathcal{B},\mathcal{B}'}(id_{\mathbb{R}^2}) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{3}{2} & \frac{1}{2} \end{pmatrix}.$$

Soluzione Esercizio 5.6.10

Come prima cosa cerchiamo una base per V; risolvendo il sistema che lo definisce, avremo:

$$\begin{cases} x_1 + x_4 = 0 \\ -2x_1 + x_3 - x_4 = 0 \\ x_2 + x_4 = 0 \end{cases} \to \begin{cases} x_1 = -x_4 \\ 2x_4 + x_3 - x_4 = 0 \\ x_2 = -x_4 \end{cases} \to \begin{cases} x_1 = -x_4 \\ x_3 = -x_4 \\ x_2 = -x_4 \end{cases}$$

Quindi i vettori di V si possono scrivere come $(-x_4, -x_4, -x_4, x_4)$: perciò una base di V è $\{(-1, -1, -1, 1)\}$, e $f(V)$ sarà generato da $f(-1, -1, -1, 1)$; per calcolare esplicitamente questo vettore, scriviamo la matrice:

$$M_{\mathcal{E}_4,\mathcal{E}_4}(f) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ -2 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

e avremo:

$$f(-1,-1,-1,1) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ -2 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

quindi $f(-1,-1,-1,1) = (0,0,0,0)$, il che ci dice che $f(V) = \{\mathbf{0}\}$ è lo spazio nullo, quindi $\dim f(V) = 0$ e $f(V)$ non ha basi.

Si può osservare che se andremo a calcolare $\ker f$ (farlo per esercizio), si otterrà che $\ker f = V$.

Soluzione Esercizio 5.6.23

Applicando la funzione ai vettori della base $\mathcal{B}$ si ha:

$$f(x^3) = 3x^2 = \begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \end{pmatrix}_{\mathcal{B}} \quad ; \quad f(x^2) = 2x = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}_{\mathcal{B}} \quad ;$$

$$f(x) = 1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}_{\mathcal{B}} \quad ; \quad f(1) = 0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}_{\mathcal{B}} \quad .$$

e quindi

$$M_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Soluzione Esercizio 5.6.30

Poiché $\ker f$, è costituito dalle soluzioni del sistema:

$$\begin{cases} 2x_1 + hx_2 + 3x_3 = 0 \\ -x_1 + 2x_2 - x_3 + hx_4 = 0 \\ x_1 + 2x_3 - 2x_4 = 0 \end{cases};$$

avremo che (riarrangiando l'ordine delle equazioni):

$$\dim \ker f = 4 - r \begin{pmatrix} 1 & 0 & 2 & -2 \\ -1 & 2 & -1 & h \\ 2 & h & 3 & 0 \end{pmatrix}$$

e, con un eliminazione di Gauss:

$$\begin{pmatrix} 1 & 0 & 2 & -2 \\ -1 & 2 & -1 & h \\ 2 & h & 3 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 2 & -2 \\ 0 & 2 & 1 & h-2 \\ 0 & h & -1 & 4 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 0 & 2 & -2 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & -1-\frac{h}{2} & -\frac{h^2-2h-8}{2} \end{pmatrix}$$

Si ha che per $h \neq -2$, il rango della matrice è 3, quindi $\dim(\ker f)$ varrà $4-3=1$. Invece per $h=-2$ il rango è 2 e quindi $\dim(\ker f) = 4-2=2$.

Ponendo $h=-2$, per determinare $\ker f$, risolvamo il sistema:

$$\begin{cases} 2x_1 - 2x_2 + 3x_3 = 0 \\ x_1 + 2x_3 - 2x_4 = 0 \end{cases} \to \begin{cases} 2(2x_4 - 2x_3) - 2x_2 + 3x_3 = 0 \\ x_1 = 2x_4 - 2x_3 \end{cases} \to$$

$$\to \begin{cases} 4x_4 - x_3 - 2x_2 = 0 \\ x_1 = 2x_4 - 2x_3 \end{cases} \to \begin{cases} x_3 = 4x_4 - 2x_2 \\ x_1 = 2x_2 - 6x_4 \end{cases};$$

quindi, dando alle variabili libere i valori $x_2 = 1, x_4 = 0$ e $x_2 = 0, x_4 = 1$, otterremo una base di $\ker f$ è data da $\left\{ \begin{pmatrix} 2 & 1 \\ -2 & 0 \end{pmatrix}, \begin{pmatrix} -6 & 0 \\ 4 & 1 \end{pmatrix} \right\}$.

Presa una base di $\mathbb{R}^{2,2}$:

$$\mathbf{v}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix};\ \mathbf{v}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ \mathbf{v}_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ \mathbf{v}_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

abbiamo che $\{f(\mathbf{v}_1), f(\mathbf{v}_2), f(\mathbf{v}_3), f(\mathbf{v}_4)\}$ è un insieme di generatori per $\mathrm{Im} f$; poiché $\dim(\mathrm{Im} f) = 4 - \dim(\ker f) = 2$, una base di $\mathrm{Im} f$ è data da due di essi che siano linearmente indipendenti; per esempio da:

$$f(\mathbf{v}_1) = \begin{pmatrix} 2 & -1 \\ -1 & 0 \end{pmatrix}, \quad f(\mathbf{v}_4) = \begin{pmatrix} 0 & -2 \\ -2 & -2 \end{pmatrix}.$$

13.6 Soluzioni Capitolo 6

Soluzione Esercizio 6.6.3

Calcoliamo il determinante delle due seguenti matrici nei due metodi diversi che conosciamo.

Sviluppiamo il determinante di $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ rispetto alla prima riga:

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 \det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} - 2 \det \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix} + 3 \det \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix} =$$
$$= 1(5 \cdot 9 - 6 \cdot 8) - 2(4 \cdot 9 - 6 \cdot 7) + 3(4 \cdot 8 - 5 \cdot 7) = -3 + 12 - 9 = 0.$$

Calcoliamo invece il determinante di $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 2 & 1 & 0 \end{pmatrix}$ utilizzando Sarrus.

$$\det \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 2 & 1 & 0 \end{pmatrix} = (1\cdot 1\cdot 0 + 0\cdot 3\cdot 2 + 2\cdot 0\cdot 1) - (1\cdot 3\cdot 1 + 0\cdot 0\cdot 0 + 2\cdot 1\cdot 2) =$$
$$= 0 - 7 = -7.$$

Soluzione Esercizio 6.6.8

Osserviamo che $\det \begin{pmatrix} k & 3 \\ 0 & k^2 \end{pmatrix} = k^3$ dunque il determinate di tale matrice sarà 0 se e solo se $k = 0$ che equivale a dire che il determinate di essa è 3 per ogni $k \neq 0$. Per $k = 0$ la matrice diventa $\begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$ che ha ovviamente rango 1.

Ora $\det \begin{pmatrix} k & 0 & -1 \\ 2-k & k & 1 \\ 0 & -3 & 1 \end{pmatrix} = k^2 + 3(2-k) + 3k = k^2 + 6$ che è diverso da 0 per ogni k reale. Dunque questa matrice è di rango massimo (ossia 3) per ogni $k \in \mathbb{R}$.

Soluzione Esercizio 6.6.18

I tre vettori $\mathbf{v}_1 = (k, 2, 1)$, $\mathbf{v}_2 = (0, 1, k)$, $\mathbf{v}_3 = (1, 3, 2)$ sono linearmente indipendenti per tutti i valori di k che renderanno la matrice da essi formata di rango 3. Osseviamo che tale matrice risulta essere $\begin{pmatrix} k & 2 & 1 \\ 0 & 1 & k \\ 1 & 3 & 2 \end{pmatrix}$ e dunque quadrata 3×3. Il che significa che è sufficiente controllare quando il suo determinante è non nullo. Ora $\det \begin{pmatrix} k & 2 & 1 \\ 0 & 1 & k \\ 1 & 3 & 2 \end{pmatrix} = -3k^2 + 4k - 1$. Ora $-3k^2 + 4k - 1 = 0$ se e solo se $k = 1/3, 1$. Dunque per $k = 1/3, 1$ i tre vettori saranno linearmente dipendenti, mentre per tutti gli altri valori di di k saranno linearmente indipendenti.

Soluzione Esercizio 6.6.22

Osserviamo che $\det \begin{pmatrix} k & 2 \\ 2 & k \end{pmatrix} = k^2 - 4 = 0$ se e solo se $k = \pm 2$. Dunque per $k \neq \pm 2$ tale matrice sarà invertibile.

Ora $\det \begin{pmatrix} -k & 0 & 0 \\ k & k & 0 \\ 1 & k & 1 \end{pmatrix} = -k^2 = 0$ se e solo se $k = 0$ dunque quest'ultima matrice sarà invertibile per ogni $k \neq 0$.

Soluzione Esercizio 6.6.31

Iniziamo osservando che se A è invertibile esiste la sua inversa che indichiamo con A^{-1}. La proprietà della matrice inversa è che

$A \cdot A^{-1} = A^{-1} \cdot A = Id$.

Dunque $\det(A \cdot A^{-1}) = \det(Id)$. Ovviamente $\det(Id) = 1$, quindi $\det(A \cdot A^{-1}) = 1$ Ora possiamo utilizzare il teorema di Binet e osservare che $\det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1})$. Dunque $\det(A) \cdot \det(A^{-1}) = 1$ che significa che sia $\det(A)$ che $\det(A^{-1})$ sono non nulli, quindi possiamo scrivere $\det(A^{-1}) = 1/\det(A)$.

Soluzione Esercizio 6.6.33

Il metodo di Kramer sappiamo che si può applicare quando la matrice dei coefficienti delle variabili è una matrice invertibile. In questo caso la matrice in questione ha 3 variabili e 2 equazioni, quindi la matrice dei coefficienti associata sarà una matrice 2×3, quindi essendo rettangolare non potrà essere invertibile in quanto la nozione di matrice inversa come l'abbiamo definita ha senso solo per matrici quadrate. Perciò il metodo di Kramer standar non può essere applicato.

Osserviamo inoltre che avendo due equazioni e 3 variabili, le soluzioni di questo sistema dipenderanno da almeno una variabile. Ha quindi senso consideare solo due delle variabili il questione (ad esempio x e y) come variabili "vere" e z "alla stregua" dei termini noti cosicché quando si calcoleranno le soluzioni esse dipenderanno in effetti da z.

La matrice dei coefficienti di x e y è $\begin{pmatrix} 1 & 1 \\ 1 & -3 \end{pmatrix}$ che effettivamente è invertibile. Applichiamo quindi Kramer utilizzando come colonna dei "termini noti" $\begin{pmatrix} z-3 \\ z+1 \end{pmatrix}$ e otteniamo

$$x = \det(\begin{pmatrix} z-3 & 1 \\ z+1 & -3 \end{pmatrix})/(-4), \quad y = \det(\begin{pmatrix} 1 & z-3 \\ 1 & z+1 \end{pmatrix})/(-4),$$

$$x = 3/4(z-3) + 1/4(z+1), \quad y = -1/4(z+1) + 1/4(z-3),$$

$$x = z - 2, \quad y = -1.$$

13.7 Soluzioni Capitolo 7

Soluzione Esercizio 7.5.5

Iniziamo con lo scrivere la matrice associata alla f rispetto alle basi canoniche:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Calcoliamo ora il polinomio caratteristico associato:

$$\det(A - \lambda Id) = \det\left(\begin{pmatrix} 1-\lambda & 0 & 1 \\ 0 & -1-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{pmatrix}\right) = (1-\lambda)^2(-1-\lambda).$$

Dunque gli autovalori di f sono $\lambda_1 = 1$ e $\lambda_2 = -1$. Osserviamo subito che dalla struttura del polinomio caratteristico si vede che la molteplicità algebrica di λ_1 è 2, mentre quella di λ_2 è 1.

Calcoliamo ora gli autovettori associati a λ_1. Dobbiamo risolvere il sistema

$$(A - Id) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Si vede semplicemente che esso ha soluzione $y = z = 0$. Dunque gli autovalori di f sono tutti del tipo $(x, 0, 0)$ con $x \in \mathbb{R}$. La cosa si esprime anche dicendo che l'autospazio V_1 associato a $\lambda_1 = 1$ è $V_1 = \{(x, y, z) \in \mathbb{R}^3 \,|\, y = z = 0\}$. Perciò λ_1 è un autovettore di molteplicità geometrica 1[1].

Calcoliamo ora gli autovettori associati a λ_2. Dobbiamo risolvere il sistema

$$(A + Id) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Si vede semplicemente che esso ha soluzione $x = z = 0$. Dunque gli autovalori di f sono tutti del tipo $(0, y, 0)$ con $y \in \mathbb{R}$. La cosa si esprime anche dicendo che l'autospazio V_{-1} associato a $\lambda_2 = -1$ è $V_{-1} = \{(x, y, z) \in \mathbb{R}^3 \,|\, x = z = 0\}$. In questo caso la molteplicità geometrica di $\lambda_2 = -1$ è 1, cosa che era già deducibile dal fatto che la sua molteplicità algebrica era 1.

Soluzione Esercizio 7.5.6

Un'applicazione lineare per essere ben definita deve essere innanzitutto una funzione; dopodiché deve soddisfare le proprietà della linearità.

Iniziamo col controllare che sia una funzione. Per prima cosa osserviamo che i vettori $\{(1,1,0), (0,1,1), (1,0,1)\}$ formano una base del dominio. Dunque la f è definita su tre vettori di base, il che significa che è ben definta su ogni vettore del dominio in quanto, per definizione di base, per ogni vettore (x, y, z) del dominio $\mathbb{R}^3$ esisteranno sempre e saranno pure unici dei parametri $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}^3$ tali che il nostro vettore si possa scrivere come combinazione lineare tramite i λ_i, $i = 1, 2, 3$, dei vettori di base: $(x, y, z) = \lambda_1(1,1,0) + \lambda_2(0,1,1) + \lambda_3(1,0,1)$. Il fatto che dei parametri λ_i siffatti esistano dimostra che ogni vettore del domino ha un'immagine. Il fatto che tali parametri siano unici dimostra che ogni vettore del dominio ha una sola immagine. Questi due fatti assieme sono sufficienti a dimostrare che f sia effettivamente una funzione.
Osserviamo che per dimostrare che f è una funzione abbiamo usato solo il fatto che è stata definita su una base del dominio.

[1] In questo esercizio non è richiesta la discussione riguardo la semplicità di f ma possiamo comunque osservare che il fatto che la molteplictà algebrica di λ_1 sia diversa da quella geometrica sarebbe sufficiente per affermare che f non è semplice.

Per quanto riguarda la linearità, se una funzione lineare con le proprietà richieste esiste dovrà avere la seguente matrice associata rispetto alle basi $\{(1,1,0),(0,1,1),(1,0,1)\}$ del dominio e la canonica del codominio:

$$A = \begin{pmatrix} 5 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

e sarà tale che $\forall\,(x,y,z) \in \mathbb{R}^3$, $f((x,y,z)) = A \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. E questo ovviamente basta per la linearità di f in quanto $\forall\,(x_1,y_1,z_1),(x_2,y_2,z_2) \in \mathbb{R}^3$ e $\forall\,\alpha_1,\alpha_2 \in \mathbb{R}$ si ha che
$f(\alpha_1(x_1,y_1,z_1) + \alpha_2(x_2,y_2,z_2)) = f(\alpha_1 x_1 + \alpha_2 x_2, \alpha_1 y_1 + \alpha_2 y_2, \alpha_1 z_1 + \alpha_2 z_2) =$

$$A \begin{pmatrix} \alpha_1 x_1 + \alpha_2 x_2 \\ \alpha_1 y_1 + \alpha_2 y_2 \\ \alpha_1 z_1 + \alpha_2 z_2 \end{pmatrix} = \alpha_1 A \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \alpha_2 A \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} =$$

$$= \alpha_1 f(x_1,y_1,z_1) + \alpha_2 f(x_2,y_2,z_2)).$$

Osserviamo che di nuovo non abbiamo utilizzato altro che la f sia stata definita su una base del domino.

Passiamo ora al calcolo di autovalori e autovettori. Ricordiamoci che per utilizzare la tecnica del polinomio caratterisitico presentata in questo capitolo occorre che la matrice che associamo alla f sia quella relativa alle basi canoniche sia nel domino che nel codominio (non la A di cui sopra). La matrice associata alla f rispetto alle basi canoniche è

$$B = \begin{pmatrix} 2 & 3 & -3 \\ 2 & 3 & -2 \\ -1 & 1 & 1 \end{pmatrix}.$$

Ora con questa matrice possiamo andare a scrivere il polinomio caratteristico

$$\det(B - \lambda Id) = \det \begin{pmatrix} 2-\lambda & 3 & -3 \\ 2 & 3-\lambda & -2 \\ -1 & 1 & 1-\lambda \end{pmatrix} =$$

$$= (\lambda - 5)(\lambda - 1/2 - \sqrt{5}/2)(\lambda - 1/2 + \sqrt{5}/2).$$

Dunque $5, 1/2 + \sqrt{5}/2, 1/2 - \sqrt{5}/2$ sono autovalori di f.

Per trovare gli autovettori di f dobbiamo risolvere i seguenti sistemi:

$$(B - 5Id) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \tag{13.1}$$

$$(B - (1/2 + \sqrt{5}/2)Id) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \tag{13.2}$$

$$(B - (1/2 - \sqrt{5}/2)Id) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \qquad (13.3)$$

Lo spazio delle soluzioni del sistema (13.1) è l'autospazio associato all'autovalore $\lambda_1 = 5$ ed è $V_{\lambda_1} = \{(x, y, z) \in \mathbb{R}^3 \,|\, x = y, z = 0\}$, quindi gli autovettori associati a $\lambda_1 = 5$ sono tutti del tipo $(x, x, 0)$ con $x \in \mathbb{R}$.

Lo spazio delle soluzioni del sistema (13.2) è l'autospazio associato all'autovalore $\lambda_2 = 1/2 + \sqrt{5}/2$ ed è
$V_{\lambda_2} = \{(x, y, z) \in \mathbb{R}^3 \,|\, x = -6z(-2 + \sqrt{5})/(1 + 2\sqrt{5})(-3 + \sqrt{5}), y = (3 + \sqrt{5})z/(1 + 2\sqrt{5})\}$,
quindi gli autovettori associati a $\lambda_2 = 1/2 + \sqrt{5}/2$ sono tutti del tipo
$(-6z(-2 + \sqrt{5})/(1 + 2\sqrt{5})(-3 + \sqrt{5}), (3 + \sqrt{5})z/(1 + 2\sqrt{5}), z)$, $z \in \mathbb{R}$.

Lo spazio delle soluzioni del sistema (13.3) è l'autospazio associato all'autovalore $\lambda_3 = 1/2 - \sqrt{5}/2$ ed è
$V_{\lambda_3} = \{(x, y, z) \in \mathbb{R}^3 \,|\, x = -6z(-2 + \sqrt{5})/(-1 + 2\sqrt{5})(3 + \sqrt{5}),$
$y = (-3 + \sqrt{5})z/(-1 + 2\sqrt{5})\}$,
quindi gli autovettori associati a $\lambda_3 = 1/2 - \sqrt{5}/2$ sono tutti del tipo
$(6z(2 + \sqrt{5})/(-1 + 2\sqrt{5})(3 + \sqrt{5}), (-3 + \sqrt{5})z/(-1 + 2\sqrt{5}), z)$, $z \in \mathbb{R}$.

Soluzione Esercizio 7.5.26

La matrice associata alla f rispetto alle basi canoniche è

$$A = \begin{pmatrix} 1 & a \\ b & 1 \end{pmatrix}.$$

Dunque il polinomio caratteristico risulta essere

$$\det(A - \lambda Id)) = \begin{pmatrix} 1 - \lambda & a \\ b & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2 - ab$$

che ha soluzioni in λ: $\lambda_{1,2} = 1 \pm \sqrt{ab}$.
Quindi se $a, b \neq 0$ e almeno uno tra a e b è negativo allora gli autovalori sono complessi e la f non è semplice.
Se $a, b > 0$ allora la f ha due autovalori distinti dunque è semplice.
Se almeno uno tra a e b è nullo allora la f ha due autovalori coincidenti. Per controllare se in questo caso è semplice occorre andare a calcolare la molteplicità geometrica di quell'unico autovalore ossia $\lambda = 1$. Ora, essendo la f un endomorfismo di $\mathbb{R}^2$ con un autovalore di molteplicità algebrica 2, sarà semplice solo se l'unico autospazio coincide con tutto $\mathbb{R}^2$, ossia se la matrice $\begin{pmatrix} 1 - \lambda & a \\ b & 1 - \lambda \end{pmatrix}$ per $\lambda = 1$ è la matrice nulla. È immediato verificare che lo è solo se sia a che b sono entrambi nulli.

Riassumento, i valori di a e b per cui la f è semplice sono: o $a, b > 0$ o $a = b = 0$.

Soluzione Esercizio 7.5.32

Il polinomio caratteristico di f è

$$\det\left(\begin{pmatrix} 1-\lambda & -k & 0 \\ 0 & 1-\lambda & -k \\ -1 & 0 & 1-\lambda \end{pmatrix}\right) =$$

$$= 1 - 3\lambda + 3\lambda^2 - \lambda^3 - k^2.$$

Dunque gli autovalori di f sono
$\lambda_1 = ((-k)^{2/3} + 1)$, $\lambda_{2,3} = -(-k)^{2/3}/2 + 1 \pm \sqrt{3}(-k)^{2/3}i/2$.

Ora se $k \neq 0$ gli autovalori λ_2 e λ_3 sono distinti e complessi, dunque in tal caso la f non potrà essere semplice in quanto avrà un solo autospazio reale (quello relativo a λ_1) di molteplicità geometrica 1 (essendo la molteplicità algebrica di λ_1 uguale a 1, lo è anche la geometrica). Se invece $k = 0$ abbiamo che $\lambda_1 = \lambda_2 = \lambda_3 = 1$, dunque abbiamo un solo autovalore di molteplicità algebrica 3. Affinché la f fosse semplice per $k = 0$ bisognerebbe che l'autospazio relativo all'autovalore 1 fosse tutto $\mathbb{R}^3$ il che equivale a richiedere che tale autospazio abbia dimensione 3, ossia che il rango della matrice $\begin{pmatrix} 1-\lambda & -k & 0 \\ 0 & 1-\lambda & -k \\ -1 & 0 & 1-\lambda \end{pmatrix}$ con $k = 0$ e $\lambda = 1$ fosse 0, cioè che tale matrice fosse la matrice nulla. Ma ciò non si verifica infatti sostituendo $k = 0$ e $\lambda = 1$ si ottiene $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$.
Quindi la f non è semplice per nessun valore di k.

13.8 Soluzioni Capitolo 8

Soluzione Esercizio 8.6.6

I vettori $\mathbf{v}$ e $\mathbf{w}$ sono ortogonali solo se il loro prodotto scalare è nullo quindi $\mathbf{v} \cdot \mathbf{w} = (1, h, 3) \cdot (h, 2, -4) = h + 2h - 12 = 3h - 12 = 0$ solo se $h = 4$. Dunque $\mathbf{v} \perp \mathbf{w}$ se e solo se $h = 4$.

Soluzione Esercizio ??

Iniziamo col trovare una base per W. È semplice vedere che essa può per esempio essere descritta dai seguenti vettori:
$\mathbf{v}_1 = (-1, 0, 1, 0), \mathbf{v}_2 = (0, 1, 0, 2)$.

Ora per trovare $W^\perp$ occorre trovare tutti i vettori ortogonali a $\mathbf{v}_1, \mathbf{v}_2$, ossia tutti i vettori $(x, y, z, t) \in \mathbb{E}^4$ tali che
$(x, y, z, t) \cdot (-1, 0, 1, 0) = 0$ e $(x, y, z, t) \cdot (0, 1, 0, 2) = 0$.

Da qui si ricavano le due seguenti equazioni: $-x + z = 0$ e $y + 2t = 0$ quindi $W^\perp = \{(x, y, z, t) \in \mathbb{E}^4 \mid -x + z = 0, y + 2t = 0\}$ quindi una sua base sarà $\{(1, 0, 1, 0), (0, 2, 0, 1)\}$.

Soluzione Esercizio 8.6.23

In ogni spazio vettoriale euclideo si ha: $\mathbf{v}_1 \cdot \mathbf{v}_2 = ||\mathbf{v}_1|| \cdot ||\mathbf{v}_2|| \cos(\mathbf{v}_1, \mathbf{v}_2)$. Ora $\mathbf{v}_1 \cdot \mathbf{v}_2 = 3 - 1 = 2$. Inoltre $||\mathbf{v}_1|| = \sqrt{2}$ e $||\mathbf{v}_2|| = \sqrt{11 + h^2}$ e $\cos(60^o) = 1/2$. Quindi abbiamo $2 = \sqrt{2}\sqrt{11 + h^2}/2$ da cui $h^2 = -3$. Perciò non esiste nessun valore di h reale tale che l'angolo tra $\mathbf{v}_1$ e $\mathbf{v}_2$ sia di 60 gradi.

Per quanto riguarda l'ortogonalità , essendo $\mathbf{v}_1 \cdot \mathbf{v}_2 = 2 \neq 0$ per ogni $h \in \mathbb{R}$ si ha che $\mathbf{v}_1$ e $\mathbf{v}_2$ non sono nemmeno ortogonali per nessun valore di h.

Soluzione Esercizio 8.6.27

Scegliamo come base ortonormale la base canonica. Rispetto a essa la matrice associata a f risulta essere:

$$A := \begin{pmatrix} 0 & h & 0 \\ -h & 0 & 0 \\ 0 & 0 & h \end{pmatrix}.$$

Affinché f sia un'isometria occorre che A sia una matrice ortogonale. Vi sono vari modi di imporlo. O cerchiamo gli h per cui $A \cdot A^t = Id$ oppure osserviamo che i tre vettori colonna di A sono banalmente ortogonali tra loro, e l'unico modo che hanno per essere tutti di norma 1 è che $h = \pm 1$. Questo è necessario e sufficiente affinché f sia un'isometria.

13.9 Soluzioni Capitolo 9

Soluzione Esercizio 9.5.2

L'equazione cartesiana di un piano π è della forma $ax + by + cz + d = 0$ dove il vettore (a, b, c) rappresenta un vettore perpendicolare a π. Dunque π sarà parallelo a π' se il vettore (a, b, c) sarà parallelo al vettore $(1, -3, 1)$ che rappresenta un vettore perpendicolare al piano π'. Dunque $(a, b, c) = \alpha(1, -3, 1)$ per un qualunque $\alpha \in \mathbb{R} \setminus \{0\}$. Scegliamo $\alpha = 1$. Perciò la condizione di parallelismo tra π e π' implica che l'equazione di π sia del tipo: $x - 3y + z + d = 0$. Imponiamo ora che $A \in \pi$: $2 + 3 - 1 + d = 0$ perciò $d = -4$. Ecco che l'equazione cartesiana di π è $x - 3y + z - 4 = 0$.

Soluzione Esercizio 9.5.5

L'equazione cartesiana di un piano π è data da

$$ax + by + cz + d = 0$$

dove il vettore (a, b, c) rappresenta un vettore perpendicolare a π.
Il piano π deve contenere i punti A e B, in particolare dovrà essere parallelo al vettore $A - B = (1, -2, -1)$. Dunque π sarà parallelo sia ad $A - B$ che a $\overrightarrow{\mathbf{v}}$. Perciò il vettore (a, b, c) dovrà essere perpendicolare sia

ad $A - B$ che a $\vec{\mathbf{v}}$. Dunque $(a, b, c) \cdot (A - B) = 0$ e $(a, b, c) \cdot \vec{\mathbf{v}} = 0$. Dalla prima equazione ricaviamo $a - 2b - c = 0$, dalla seconda $a + b = 0$. Perciò: $a = -b$ e $c = -3b$. Ecco che il vettore (a, b, c) può essere scritto nella forma $(-b, b, -3b)$. Ovviamente per qualunque valore di $b \neq 0$ tale vettore sarà perpendicolare a π, scegliamo $b = 1$ e otteniamo π: $-x + y - 3z + d = 0$. Per ricavare d imponiamo il passaggio per uno qualunque dei due punti A o B. Così facendo troviamo $d = 6$ perciò π: $-x + y - 3z + 6 = 0$.

Soluzione Esercizio 9.5.6

Il piano π di equazione $ax + by + cz + d = 0$ per essere parallelo alla retta retta r, dovrà essere tale che il vettore (a, b, c) (che è perpendicolare a π) sia perpendicolare al vettore direttore della retta r (in quanto il vettore direttore di una retta è parallelo alla retta stessa). Dunque $(a, b, c) \cdot (1, 2, 3) = 0$ da cui $a + 2b + 3c = 0$, quindi $a = -2b - 2c$. Il piano π deve inoltre contenere due punti, perciò sarà parallelo al vettore direttore della retta da essi individuata: quindi di nuovo (a, b, c) dovrà essere perpendicolare al vettore $P - Q = (-1, 0, -1)$, perciò $(-2b - 2c, b, c) \cdot (-1, 0, -1) = 0$, da cui $2b + 2c - c = 0 \Rightarrow c = -2b$, dunque $(a, b, c) = (-2b, b, -2b)$. Ecco che π diventa $-2x + y - 2z + d$. Per trovare d imponiamo il passaggio di π da P e otteniamo $-2 + 1 - 2 + d = 0$, quindi $d = 3$, e così π diventa $-2x + y - 2z + 3 = 0$.

Soluzione Esercizio 9.5.13

La retta r scritta in forma parametrica risulta:

$$r : \begin{cases} x = lt + x_0 \\ y = mt + y_0 \\ z = nt + z_0 \end{cases}.$$

Il passaggio per il punto Q impone semplicemente che diventi:

$$r : \begin{cases} x = lt + 1 \\ y = mt - 1 \\ z = nt \end{cases}.$$

Il fatto che r sia parallela a π equivale a $(l, m, n) \cdot (2, -1, 0) = 0$ e quindi $m = 2l$. Perciò il vettore (l, m, n) diventa $(l, 2l, n)$.
Qualunque valore si scelga per l, n ci darà una retta della forma cercata.

Soluzione Esercizio 9.5.15

Se cerchiamo le equazioni parametriche della retta, occorre conoscere un punto per cui passa la retta (e possiamo scegliere indifferentemente o A o B) e un vettore che ne indichi la direzione. Per quest'ultimo basta prendere il vettore $A - B = (1, 2, 3) - (0, -4, 1) = (1, 6, 2)$. Quindi le equazioni parametriche della retta sono:

$$\begin{cases} x = t + 1 \\ y = 6t + 2 \\ z = 2t + 3 \end{cases}.$$

Soluzione Esercizio 9.5.16

Se cerchiamo le equazioni parametriche della retta, occorre conoscere un punto per cui passa la retta (P nel nostro caso) e un vettore che ne indichi la direzione. Per quest'ultimo dobbiamo trovare un vettore ortogonale a π. Il piano ci è dato in forma parametrica, quindi abbiamo i vettori che individuano la direzione del piano: $(1,0,1)$ e $(0,1,0)$ (i coefficienti di t_1 e t_2 rispettivamente). Per avere un vettore perpendicolare a π basta imporre la condizione di ortogonalità a questi ultimi due vettori: cerchiamo il vettore (x,y,z) tale che $(x,y,z)\cdot(1,0,1) = (x,y,z)\cdot(0,1,0) = 0$, da cui

$$\begin{cases} x+z=0 \\ y=0 \end{cases}.$$

Perciò un vettore direttore della retta sarà $(-1,0,1)$, quindi le equazioni cercate sono:

$$\begin{cases} x=-t+1 \\ y=0 \\ z=t \end{cases}.$$

Soluzione Esercizio 9.5.17

Affiché la retta cercata sia parallela a π occorre che il suo vettore direttore sia una combinazione lineare dei vettori direttori di π, quindi il vettore direttore della retta cercata sarà

$a(-1,0,-1)+b(1,3,-2) = (b-a,3b,-a-2b)$.

Ora dobbiamo imporre che essa sia pure sghemba con r. Per fare questo prediamo il vettore direttore appena trovato $(b-a,3b,-a-2b)$, quello della retta r ossia $(1,-1,-2)$ e un terzo vettore ottenuto come differenza tra due punti qualunque della retta cercata ed r: come punto della retta cercata prendiamo B, come punto di r prendiamo quello che si ottiene per $t=0$, quindi questo vettore congiungente le due rette sarà

$(3,1,-2)-(3,0,-2) = (0,1,0)$.

Ora, affinchè le due rette siano sghembe occorre che i tre vettori di cui sopra non siano complanari, quindi basta che siano linearmente indipendenti[2]. Imponiamo quindi la condizione di lineare indipendenza:

$$\begin{vmatrix} b-a & 3b & -a-2b \\ 1 & 1 & -2 \\ 0 & 1 & 0 \end{vmatrix} \neq 0,$$

quindi $-2(b-a)-(-a-2b)\neq 0$, ossia $a\neq 0$. Per qualunque valore di b e qualunque valore non nullo di a avremo la retta cercata. Prendiamo per semplicità $a=1, b=0$: $(-1,0,-1)$, quindi la retta cercata avrà equazioni parametriche:

$$\begin{cases} x=-t+3 \\ y=1 \\ z=-t-2 \end{cases}.$$

[2]Osserviamo che questa è quella che si suole chiamare una condizione aperta, quindi vi sarà un'infinità di vettori a soddisfare questa condizione, ma questo non deve stupire in quanto vi è un'infinità di rette passanti per un punto, parallele a un piano e sghembe con una retta.

Soluzione Esercizio 9.5.18

Per trovare la posizione reciproca delle due rette studiamo la matrice dei coefficienti associata al sistema che individua la loro intersezione:

$$\left(\begin{array}{ccc|c} 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & -k & 1 \\ 0 & 1 & 0 & 0 \end{array}\right).$$

Osserviamo che il rango della matrice incompleta è 3 indipendentemente dal valore di k (le tre righe che non coinvolgono k rappresentano ovviamente 3 vettori linearmente indipendenti). Questo significa che le due rette potranno essere solo o incidenti (se il rango della matrice completa sarà 3) o sghembe (se il rango della matrice completa sarà 4). Ora il determinate della matrice completa è 0 per ogni valore di k, quindi le due rette date sono incidenti per ogni valore di k.

13.10 Soluzioni Capitolo 10

Soluzione Esercizio 10.5.3 La matrice associata a φ rispetto alla base canonica è

$$M_{\mathcal{E}}(\varphi) = \begin{pmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

che è una matrice simmetrica a determinante nullo quindi φ è una forma bilineare simmetrica degenere. I vettori isotropi di φ sono quelli per cui $\varphi(v,v) = 0$; sia quindi $v = (x,y,z) \in \mathbb{R}^3$, $\varphi(v,v) = x^2 + 2xy - y^2 = 0$ solo se $(x+y)^2 - 2y^2 = 0$ ossia se $(x + (1+\sqrt{2})y)(x + (1-\sqrt{2}y)) = 0$ quindi i vettori isotropi sono tutti della forma $(-(1+\sqrt{2})y, y, z)$ oppure $(-(1-\sqrt{2})y, y, z)$.

Soluzione Esercizio 10.5.4 La matrice associata a φ rispetto alla base canonica è

$$M_{\mathcal{E}}(\varphi) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

i cui autovalori sono $\lambda = \pm 1$ quindi la sua forma di Sylvester è

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

e φ non è un prodotto scalare.

Soluzione Esercizio 10.5.9 Nessuna delle tre quadriche è a centro.

13.11 Soluzioni Capitolo 11

Soluzione Esercizio 11.4.1

Il sistema:

$$\begin{cases} 3x+2y-z-t=1 \\ x+y+z+t=2 \\ 2x+y-2z-2t=-1 \end{cases}$$

è del tipo $AX = B$, ove $A = \begin{pmatrix} 3 & 2 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & -2 & -2 \end{pmatrix}$; poiché si ha che $r(A) = 2$ (la terza riga si ottiene dalla prima meno la seconda), allora si ha $\dim W = 4 - 2 = 2$, ove $W \subset \mathbb{R}^4$ è il sottospazio vettoriale associato al sottospazio affine (W è dato dalle soluzioni di $AX = O$), quindi anche la dimensione del sottospazio affine è 2.

Soluzione Esercizio 11.4.4

Il sottospazio affine $S_1 = \overline{P_0, P_1} \subset \mathbb{A}^4$ è parallelo al sottospazio vettoriale $W \subset \mathbb{R}^4$ generato da $\mathbf{v} = \overrightarrow{P_0P_1} = (1, -1, -1, 1)$ e ha equazioni parametriche:

$$\begin{cases} x_1 = t_1 \\ x_2 = -t_1 + 1 \\ x_3 = -t_1 + 1 \\ x_4 = t_1 \end{cases}.$$

La sua intersezione con S_2 sarà data dalle soluzioni di:
$t_1 - t_1 + 1 - t_1 + 1 - t_1 = 2$, e cioè $t_1 = 0$;

quindi $S_1 \cap S_2 = P_0$.

Soluzione Esercizio 11.4.9

Il sottospazio affine S definito da $AX = B$ è parallelo al sottospazio vettoriale $W \subset \mathbb{R}^n$ definito da $AX = O$, e (per il teorema di Rouché–Capelli), $\dim W = n - r(A)$, quindi anche $\dim S = n - r(A)$, il che vale a dire che la sua codimensione è $r(A)$.

Bibliografia

Libri di riferimento per la materia, testi consigliati per esercizi, siti utili.

Testi di riferimento

C.Ciliberto: *Algebra Lineare*. Bollati Boringhieri, Torino, 1994.

S.Lang: *Algebra Lineare*. Bollati Boringhieri, Torino, 1970.

E.Sernesi: *Geometria 1*. Bollati Boringhieri, Torino, 1° ed. 1989.

K. Splinder: *Abstract Algebra with Applications* (2 vol.). M. Dekker Ed. New York, Basel, Hong Kong, 1994.

Testi di interesse particolare

M. Argeri, F. Calio, A. Lazzari, D. Sesana: *Geometria vettoriale per la grafica*. CittaStudi Ed. Novara, 2011.

G. Polya: *Come risolvere i problemi di matematica*. UTET Università , Torino, 2016.

V. Romano: *Metodi matematici per i corsi di ingegneria*. CittaStudi Ed. Novara, 2018.

Testi di esercizi

M.Abate, C. de Fabritiis: *Esercizi di Geometria*. McGraw-Hill, Milano, 1999.

A.Barani, L.Grasselli, C.Landi: *Algebra Lineare e Geometria. Quiz e Esercizi*. Progetto Leonardo, Ed. Esculapio, Bologna, 2014.

E.Carlini, M.V.Catalisano, F.Odetti, A.Oneto, M.E.Serpico: *Geometria per Ingegneria. Una raccolta di temi d'esame risolti*. Progetto Leonardo, Ed. Esculapio, Bologna, 2012.

Alcuni siti utili

M.Ferri: *CHICKEN'S CORNER: I vostri errori più frequenti*
`http://www.dm.unibo.it/ ferri/` hm/chicken.htm

G.Ottaviani: *Appunti per il corso di Geometria I e II*
`http://web.math.unifi.it/users/ottavian/index3.html`

Progetto Matematica (a cura di D.Aliffi, A.Gimigliano): *Argomenti di Matematica per i primi anni delle facoltà scientifiche*,
`http://progettomatematica.dm.unibo.it/`

Indice analitico

Affinità, 267
Anello, 22
Anello commutativo, 22
Angolo tra vettori, 177
Applicazione lineare, 96
Applicazione lineare semplice, 161
Argomento, 278
Asse, 243
Autospazio, 151, 163
Autovalore, 150, 153
Autovettore, 150, 153

Base, 33, 36, 37
Base canonica, 33
Base ortogonale, 180, 186
Base ortonormale, 180, 181, 195
Binet teorema di, 133, 254

Campo, 22
Cauchy-Schwarts disuguaglianza di, 176
Cilindro ellittico, 257, 258
Cilindro ellittico immaginario, 257
Cilindro iperbolico, 257, 258
Cilindro parabolico, 257, 258
Circonferenza, 243
Combinazione lineare, 30
Complemento algebrico, 127
Conica a centro, 254
Coniche, 245, 249, 250, 252
Coniche a centro, 254
Coniche degeneri, 249, 254, 255
Coniche non degeneri, 254
Coniugato, 274
Cono ellittico, 257, 258
Cono ellittico immaginario, 257
Coordinate affini, 263
Coordinate parametriche, 266
Coordinate polari, 277
Cramer regola di, 143

De Moivre terorema di, 279
Determinante, 127, 137, 140
Dimensione, 38
Direttrice, 243, 245, 246
Direzione di un vettore, 25
Discriminante, 252
Distanza, 177, 197, 216
Disuguaglianza triangolare, 176

Eccentricità, 245
Elemento neutro, 21, 276
Ellisse, 244, 246, 248, 250, 251, 254, 255
Ellissoide, 256, 258
Ellissoide non reale, 257
Endomorfismo, 96
Endomorfismo semplice, 161
Equazione caratteristica, 154
Esponenziale complesso, 283
Eulero formula di, 283

Fasci di piani, 218
Forma bilineare, 229
Forma bilineare non degenere, 233
Forma bilineare simmetrica , 231
Forma canonica, 238, 246, 248, 249, 255, 256
Forma normale, 241
Forma quadratica (semi)definita positiva (negativa), 238
Forma quadratica non degenere, 237
Forme quadratiche, 236
Funzione, 18
Funzione biettiva, 20
Funzione iniettiva, 20
Funzione inversa, 105
Funzione suriettiva, 20

Fuoco, 245, 246

Gauss eliminazione di, 77, 79, 135
Generatori, 30
Giacitura, 199
Gram-Schmidt ortogonalizzazione di, 181
Grassmann formula, 42
Gruppo commutativo, 21

Immagine, 99, 102, 110
Insieme, 12
Intersezione insiemistica, 14
Inverso, 21, 276
Iperbole, 244, 246, 248, 250, 251, 254, 255
Iperboloide a due falde, 257, 258
Iperboloide a una falda, 256, 258
Iperboloide ellittico, 257, 258
Iperboloide iperbolico, 256, 258
Iperpiano, 264
Isometria, 184, 186, 253
Isomorfismo, 104

Mappa esponenziale, 282
Matrice, 49
Matrice associata a una forma quadratica, 237
Matrice identità, 62
Matrice inversa, 63, 142
Matrice invertibile, 153
Matrice orlata, 140
Matrice trasposta, 54
Matrici antisimmetriche, 54
Matrici associate ad una forma bilineare, 230
Matrici del cambiamento di base, 112
Matrici di applicazioni lineari, 108
Matrici diagonali, 54
Matrici diagonalizzabili, 160
Matrici invertibili, 63, 115
Matrici ortogonali, 186
Matrici ortogonalmente diagonalizzabili, 188
Matrici quadrate, 53, 61
Matrici simili, 117
Matrici simmetriche, 54, 188, 189
Matrici triangolari, 53
Metodo del completamento dei quadrati, 239
Minore, 126, 139
Modulo, 175, 274, 278
Modulo di un vettore, 25
Molteplicità algebrica, 159
Molteplicità geometrica, 159, 164
Moltiplicazione per scalari, 26

Norma, 175
Nucleo, 99, 102, 110
Numeri complessi, 272
Numeri interi, 15
Numeri naturali, 15
Numeri razionali, 15
Numeri reali, 15

Parabola, 244, 246, 250, 251, 254, 255
Paraboloide a sella, 256, 257
Paraboloide ellittico, 256, 257
Paraboloide iperbolico, 256, 257
Parametri direttori, 209, 212
Parte immaginaria, 272
Parte reale, 272
Permutazione, 136
Permutazione dispari, 137
Permutazione pari, 137
Piani nello spazio equazione cartesiana, 198, 200, 201
Piani nello spazio equazioni parametriche, 199
Piani nello spazio incidenti, 202
Piani nello spazio paralleli, 201, 202
Piani nello spazio perpendicolari, 203
Piano affine, 264
Pivot, 80
Polinomio caratteristico, 154
Principio di induzione, 16
Prodotto cartesiano tra insiemi, 14
Prodotto scalare, 171, 173, 181, 196, 233
Prodotto scalare standard, 174
Prodotto tra matrici, 57, 58, 64
Prodotto vettoriale, 220
Proprietà associativa, 21

Proprietà commutativa, 21

Quadriche, 255, 256
Quadriche a centro, 256, 258
Quadriche degeneri, 256, 258
Quadriche non a centro, 256, 257
Quadriche non degeneri, 256

Radicale, 232
Radici n-esime, 280
Rango, 80, 84, 135, 139, 140, 253
Regola dei segni di Cartesio, 241
Retta affine, 264
Rette e piani incidenti, 213, 214
Rette e piani paralleli, 213, 214, 217
Rette e piani perpendicolari, 215
Rette nello spazio complanari, 211
Rette nello spazio equazioni cartesiane, 204, 208, 210
Rette nello spazio equazioni parametriche, 205, 207, 209, 210
Rette nello spazio parallele, 211, 212
Rette nello spazio perpendicolari, 211, 212
Rette nello spazio sghembe, 211
Rouché-Capelli teorema di, 82

Sarrus regola di, 130
Sezioni coniche, 243
Sistema lineare, 72
Sistemi lineari omogenei, 83
Somma diretta, 43
Somma tra vettori, 25
Sottospazio affine, 263
Sottospazio vettoriale, 39
Sottrazione insiemistica, 14
Spazi affini standard, 263
Spazio affine, 261
Spazio ortogonale, 178, 183
Spazio somma, 42, 162
Spazio vettoriale, 29
Spazio vettoriale euclideo, 174, 175, 181
Struttura algebrica, 21

Teorema di Gauss, Lagrange, 240, 254
Teorema di Sylvester, 234
Teorema fondamentale dell'algebra, 282
Teorema spettrale, 189, 255
Traslazioni, 267

Unione insiemistica, 14
Unità immaginaria, 272

Verso di un vettore, 25
Versori, 180
Vettore, 25, 29
Vettore nullo, 25
Vettori isotropi, 232, 237
Vettori linearmente dipendenti, 31
Vettori linearmente indipendenti, 31, 38
Vettori ortogonali, 177, 189, 232
Vettori ortonormali, 180, 181
Vettori perpendicolari, 177